Note: Text formatted in standard font is intended to be read aloud by the facilitator.
Text formatted in bold is excerpted directly from the presentation slides.
Text formatted in italics is intended as directions or notes for the facilitator; italicized text is not meant to be read aloud. Additional information can be found in the NCRTI Facilitator’s Guide which can be accessed at www.rti4success.org

Welcome the participants, to the training on Things to Consider When Developing an RTI Evaluation Plan. This training was developed by the National Center on Response to Intervention, and focuses on areas to consider when developing an RTI evaluation plan.

Audience: This presentation is for schools and districts that are implementing RTI in some capacity and where there are district- or school-level teams working to develop an RTI evaluation plan.

Length: The presentation was developed as a 4–6 hour training, but the length of the training can be adapted by changing the number of examples and the length of time provided for the activities.

Activities: There are four handouts that are embedded in the presentation. They include the Implementing the Essential Components Handout (used for note taking purposes), Measuring Fidelity Handout (5–10 minutes), the Evaluation Plan Graphic Organizer (20–30 minutes total, each section about 5 minutes), and the Guiding Questions for Evaluation Plan Development Handout (5–10 minutes). The handouts can be found at the back of the manual. Additional activities, including think-pair-shares, group questions, and reflection, are also provided in the presentation.

Note: You may consider including information from some of the other training modules that NCRTI has developed. These modules can be found at www.rti4success.org. You also may consider training only on evaluation tools that sites are expected to use, and use data from sites if available.
Review contents of the slide.

The agenda may be changed to fit the timeframe and focus of the training.
The objectives for this training are to—

*Review contents of the slide.*
First we will discuss the rationale for evaluating your RTI model and the key questions (research questions) that you hope to answer through the evaluation. Although there are many reasons to implement RTI and conduct an evaluation, it is important to have a clear sense of the purpose for RTI in order to develop a manageable RTI evaluation. Trying to answer too many questions can make the evaluation too complicated and overwhelming.
The motivation for conducting an evaluation of your school, district, or state's RTI implementation will depend on the intended purpose for RTI and the amount of time you have been implementing RTI. In looking at why you evaluate RTI, you may think about different outcomes that can tell you whether your RTI model is achieving the intended purpose. Teams should also consider whether the RTI evaluation can achieve its purpose in the given timeframe and with the available resources without placing unreasonable burdens on implementation sites. Here are just a few examples of the types of outcomes that teams may decide to look at through an evaluation of RTI.

Review contents of the slide.

Note: Depending on the stage of implementation that participants are in, it may be important to emphasize different things. For example, if the schools and districts are new to implementing RTI you should focus on the need for looking at the structures that are in place and not how well those structures are doing (that is or should be part of how the RTI model is used in practice) and then moving towards evaluating outcomes. For participants that have been implementing RTI for a number of years and have all components in place, it may be more important to focus on the effectiveness of the model by looking at student outcomes. From implementation science we know that it takes 2–4 years for full implementation.
Review slide. This is continued from the previous slide.

Note: If participants have participated in the training focused on developing technical assistance resources, consider referencing that module, or if there are questions, consider including information or referring participants to that module here. That module can be found at www.rti4success.org.
Key Evaluation Questions

1. How do I know whether my RTI model and the components of RTI are being implemented with fidelity?
2. How do I know if my RTI model is working and to what extent is it working?
3. If my RTI model is working, what will change for students, for schools, for districts?

Evaluation questions focus on two key areas. First, we need to consider whether the RTI model is being implemented and whether it is being implemented with fidelity. This includes both the implementation of the essential components (screening, progress monitoring, multi-level prevention system, and data-based decision making) and the process. If not all of the components are being implemented or the model is not being implemented with fidelity, it is unclear whether the results achieved are based on implementation of RTI or other, extraneous factors.

Second, we are interested in the outcomes of the RTI model itself, that is, to what extent is my RTI model working? If it is working, what will change for students, schools, and districts?

During the presentation today, we first will talk about the components of an RTI model and whether they are being implemented. Next we will look at how we can assess whether your RTI model and its components are being implemented with fidelity. Then we will look at some ways that you will be able to see whether the implementation of the RTI model is resulting in the intended student outcomes.
WHAT IS RTI?

One of the first steps of evaluating an RTI model is looking to see if the essential components are in place (this is or should be part of how the RTI model is used in practice). In this section we are going to provide an overview of the NCRTI definition of RTI and the essential components that make up an RTI model.

Key Terms:
- Essential components
- Screening
- Progress monitoring
- Data-based decision making
- Culturally responsive
- Evidence based

Main Points:
- RTI is a school-wide, multi-level prevention system that integrates assessment and intervention.
- RTI is preventive, not prereferral. The primary purpose of RTI is to prevent poor learning outcomes for all students. RTI may be a part of a determination process for identifying students with specific learning disabilities or other disabilities.
- The four essential components of RTI are screening, progress monitoring, the multi-level prevention system, and data-based decision making.
Defining RTI

- Response to Intervention (RTI) integrates assessment and intervention within a school-wide, multi-level prevention system to maximize student achievement and reduce behavior problems.

Although not required, it is recommended that participants have access to the one-page What Is RTI? placemat (http://www.rti4success.org/pdf/What_is_RTI_2010_07_14_placemat.pdf), a supplement to the Essential Components of RTI – A Closer Look at Response to Intervention, for easy reference.

The National Center on RTI uses a definition for RTI that includes what the Center considers to be the essential components. Response to Intervention integrates assessment and intervention within a school-wide, multi-level prevention system to maximize student achievement and reduce behavior problems. It is important to point out that RTI is a school-wide prevention system, as opposed to a prereferral process for special education, and it is multi-level as opposed to multitier. There are three levels of prevention in an RTI framework, and states, school districts, and schools can have multiple tiers within those three levels of instruction to prevent poor learning outcomes.
The second part of the definition highlights the essential components of an RTI framework.

- The first component involves **schools identifying students at risk for poor learning outcomes**. This process is commonly referred to as universal screening.
- The next component involves **monitoring student progress** through progress monitoring.
- The third component relates to **providing evidence-based interventions based on a student’s responsiveness**. It is not merely the delivery of interventions that is important; there must be a multi-level prevention system in which students have access to increasingly intense levels of instruction and interventions.
- The last component involves using data (e.g., screening or progress monitoring) to **adjust the intensity and nature of those interventions based on student responsiveness**. In other words, there is an explicit, systematic process for data-based decision making.

Some people mistakenly believe that RTI involves only special education. It is important to remember that RTI is a school-wide, multi-level prevention system that results in data that may be used as part of the determination process for identifying students with specific learning disabilities or other disabilities in accordance with your state’s law.
RTI as a Preventive Framework

- RTI is a multi-level instructional framework aimed at improving outcomes for all students.
- RTI is preventive, and provides immediate support to students who are at risk for poor learning outcomes.
- RTI may be a component of a comprehensive evaluation for students with learning disabilities.

In summary, RTI is a preventive framework. RTI is not a new name for a prereferral process. The intent of RTI is to improve outcomes for all students while providing immediate supplemental supports to students at risk for poor learning outcomes. RTI may be a component of a comprehensive evaluation for specific learning disability determination, but this is not its overarching purpose.
## Essential Components of RTI

- Screening
- Progress Monitoring
- School-Wide, Multi-Level Prevention System
  - Primary level
  - Secondary level
  - Tertiary level
- Data-based decision making for
  - Instruction
  - Evaluating effectiveness
  - Movement within the multi-level system
  - Disability identification (in accordance with state law)

### Note:
Provide a brief overview of this slide. Each component will be discussed in more detail in the upcoming slides.

So, as you saw in the definition, the Center has identified four essential components for RTI:

- **Screening**: a system for identifying students at risk for poor learning outcomes
- **Progress monitoring**: a system for monitoring the effectiveness of the supports provided to students
- **School-wide, multi-level prevention system**: at least three increasingly intense levels of instructional support
  - **Primary** – the core instruction and curriculum.
  - **Secondary** – instruction that is supplemental to the primary level that provides supports targeted to students’ needs
  - **Tertiary** – instruction that is also supplemental to primary, but more intense than secondary
- **Data-based decision making for**
  - **Instruction** – determining who needs assistance, what type of instruction or assistance is needed, whether the duration and intensity are sufficient, etc.
  - **Evaluating effectiveness** – evaluating the effectiveness of the core curriculum and instruction for all students, interventions, and the RTI framework.
  - **Movement within the multi-level system** – when to move students to something more or less intense, who is responding or not responding, etc.
  - **Disability identification** – when to refer for special education evaluation, how the student compares to his/her peers, whether he/she has received appropriate instruction, etc. This is, of course, in accordance with state law.

When evaluating the RTI model, it is important to consider whether you are implementing these components, whether you are fully implementing them, and whether you are implementing them with fidelity. In the next few slides we will cover each of the components in more detail.
The Center has developed this graphic to highlight the RTI framework. Many of you probably associate the red, green, and yellow triangle with RTI. In reality, the triangle does not represent the RTI framework; it only represents one component, the multi-level prevention system. The Center graphic takes into account all of the essential components, and most importantly, the use of data to make decisions, which is often absent from the traditional RTI triangle.

If you look to the far left, you see screening; to the far right, progress monitoring; and at the bottom, the multi-level prevention system. The three outer components require and are necessary parts of data-based decision making, which is why the arrows travel in both directions. If the three other components are in place, but data-based decision making is absent, then RTI is technically not being implemented.

In the inner ring, you will see the phrase “culturally responsive,” meaning the screening tools, progress monitoring tools, core instruction, interventions, and data-based decision making procedures should all be culturally responsive. In the same ring, you will notice the phrase “evidence-based,” implying that all components are evidence based. If these components are implemented through a cohesive model, student outcomes should improve. I’m now going to talk about each essential component in more detail.
Screening

- **PURPOSE:** Identify students who are at risk of poor learning outcomes
- **FOCUS:** All students
- **TOOLS:** Brief assessments that are valid, are reliable, and demonstrate diagnostic accuracy for predicting learning or behavioral problems
- **TIMEFRAME:** Administered more than one time per year (e.g., fall, winter, spring)

The **purpose of screening** is to identify those students who are at risk for poor learning outcomes. Because RTI is a framework for providing services, the outcomes you are concerned about could vary, and include things such as academic achievement, behavior, graduation, and post-school outcomes. Sites (state, district, schools) typically identify what outcomes students are expected to achieve and then screen to see which students are not likely to achieve those outcomes. Screening can answer these questions:

- Is our core curriculum and instruction effective?
- Which students need additional assessment and instruction?

For example, if the desired outcome is graduation, a quick screen of attendance and credits—predictors of graduation—can reveal which students are not likely to meet the requirements of graduation and will need additional support. If the desired outcome is mastery on end-of-year tests, student performance measures like curriculum-based measurements (CBMs) can reveal which students are not likely to pass the test and need additional support.

The **focus** is on all students, not just those students we may believe are at risk. Students may slip through the cracks unless there is a systematic process for screening in place. Screening is not a diagnostic test; it is a brief, reliable, and valid assessment to identify which students may need additional assessments, such as progress monitoring or diagnostic assessments, or additional instructional support. The tools should demonstrate diagnostic accuracy for predicting learning or behavioral problems. In other words, they should be able to accurately identify who could be at risk.

At a minimum, screening should be administered more than once per year, such as at the beginning of the school year and the middle of the school year. Schools and districts that wish to use screening data to evaluate program effectiveness, to establish local norms and cut scores, and to provide data to the next-year teacher typically choose to administer the screening assessment three times a year (e.g., fall, winter, spring), and should select a screening tool that provides alternative forms and multiple benchmarks.
THINK-PAIR-SHARE

- How are you implementing screening in your RTI model?

Optional: throughout this section, refer participants to the Implementing the Essential Components Handout for note taking.

Think-Pair-Share
Think about how you are implementing screening in your RTI model.

Give participants approximately 20 seconds.

Pair and share with your neighbor/table and jot down your answer.

Give participants approximately 2–3 minutes.

Allow two or three pairs/tables to orally share their answers.

What questions do you have specific to screening?

Remind participants that additional information on screening can be found in the RTI Implementer Series Module 1: Screening, which can be found at http://www.rti4success.org/resourcetype/rti-implementer-series-module-1-screening.
The purpose of progress monitoring is to monitor students’ responses to primary, secondary, or tertiary instruction. Progress monitoring data can be used to 1) estimate the rates of improvement, which allows for comparison to peers; 2) identify students who are not demonstrating adequate progress so that instructional changes can be made; and 3) compare the efficacy of different forms of instruction (in other words, identify the instructional approach or the intervention that has led to the greatest growth among students). It answers these questions:

- Are students meeting short- and long-term performance goals?
- Are students making progress at an acceptable rate?
- Does the instruction need to be adjusted or changed?

Progress monitoring is not just for those students identified for supplemental instruction. The focus is on students who have been identified through screening as being at risk for poor learning outcomes. This could include students just above the cut score as well as those scoring below the cutoff score.

Progress monitoring tools, just like screening tools, should be brief, valid, reliable, and evidence based. Common progress monitoring tools include general outcome measurements, including CBMs and mastery measurements.

The timeframe for progress monitoring assessment is dependent on the tools being used and the typical rate of growth for a student. Progress monitoring can be used any time throughout the school year. With progress monitoring, students are assessed at regular intervals (e.g., weekly, biweekly, or monthly) to produce accurate and meaningful results that teachers can use to quantify short- and long-term student progress toward end-of-year goals. At a minimum, progress monitoring tools should be administered at least monthly. However, more frequent data collection is recommended given the amount of data needed for making decisions with confidence (six to nine data points for most tools). With progress monitoring, teachers establish long-term (i.e., end-of-year) goals that indicate the level of proficiency students should demonstrate by the end of the school year.
THINK-PAIR-SHARE

- How are you implementing progress monitoring in your RTI model?

Remind participants that they can use the Implementing the Essential Components Handout for note taking.

Think-Pair-Share

**Think** about how you are implementing progress monitoring in your RTI model.

*Give participants approximately 20 seconds.*

**Pair** and **share** with your neighbor/table and jot down your answer.

*Give participants approximately 2–3 minutes.*

**Allow two or three pairs/tables to orally share their answers.**

What questions do you have specific to progress monitoring?

Within the **school-wide, multi-level prevention system**, school-wide means preventive instruction for all students, not just a series of interventions for some students. This instructional system is also designed to be **preventive**, meaning instructional supports are put in place before a student fails. It is important to remember that RTI is not trying to prevent special education. Instead, special education is another level of support designed to prevent general school failure. The RTI framework has three levels: primary, secondary, and tertiary.

The first level, or **primary level**, is indicated in green. It is expected that most students, at least 80%, should benefit from differentiated instruction within the core curriculum.

The next level, or **secondary level**, is supplemental to the primary level. It is expected that about 10–15% of students will need supplemental, small-group instruction to benefit from core instruction and curriculum.

The top level, or **tertiary level**, includes specialized, individualized instruction for students with intensive needs. It typically involves small-group instruction of 1–3 students who are significantly behind their peers. It is expected that about 5% of students will need intensive support.

If fewer than 80% of students are benefiting from the primary prevention system, consider focusing school improvement efforts on improving core instruction and curriculum. If there is a large percentage of students in the secondary or tertiary level, consider implementing large-group instructional activities and system changes with the primary level in order to reduce the number of students requiring additional support.
Data should guide decisions about changing the level of support needed for students to be successful. This change can either be an increase or decrease in the intensity of the instruction. In cases where students are responding, teams may consider decreasing the intensity. In cases where students are not responding or making adequate progress, the team may consider increasing the intensity.

There are five main approaches to changing the intensity of an intervention.

- The first is to change the intervention itself; in cases where the current intervention is believed to be ineffective for the student, the team may consider selecting a different intervention. However, this may not always be necessary. If student data indicate a student is making some progress but not necessarily adequate progress, the data-based decision-making team may decide to change the intensity of support by manipulating one or more factors of the intervention.
- Another way to address the intensity is to increase or decrease the duration of the intervention, or how long the student is receiving the intervention each time.
- The intensity may also be changed by increasing or decreasing the number of times a student participates in the intervention—the frequency of the intervention (e.g., from three to five times a week, or additional times each day).
- In some cases, the intensity may be modified by changing the interventionist. For example, some schools use paraeducators to deliver supplemental interventions. The intensity of the intervention may be changed by using a content specialist, such as a reading coach, to deliver the intervention.
- Another way to modify the intensity is to increase or decrease the number of students participating in the intervention. For example, the team may consider reducing the group size from five to two in order to provide the students more direct instruction and opportunities to respond.
THINK-PAIR-SHARE

- What does your multi-level prevention system look like? Are you fully implementing all three levels?

Remind participants that they can use the Implementing the Essential Components Handout for note taking.

Think-Pair-Share

Think about what your multi-level prevention system looks like and whether you are fully implementing all three levels.

Give participants approximately 20 seconds.

Pair and share with your neighbor/table and jot down your answer.

Give participants approximately 2–3 minutes.

Allow two or three pairs/tables to orally share their answers.

Are there any questions about the multi-level prevention system?

Remind participants that additional information on the Multi-Level Prevention System can be found in the RTI Implementer Series Module 3: Multi-Level Prevention System. The module is located at http://www.rti4success.org/resourcetype/rti-implementer-series-module-3-multi-level-prevention-system.
Data-Based Decision Making: The Basics

- Analyze data at all levels of RTI implementation (e.g., state, district, school, grade level) and all levels of prevention (i.e., primary, secondary, or tertiary).
- Establish routines and procedures for making decisions.
- Set explicit decision rules for assessing student progress (e.g., state and district benchmarks, level, and/or rate).
- Use data to compare and contrast the adequacy of the core curriculum and the effectiveness of different instructional and behavioral strategies.

Implementation of screening, progress monitoring, and the multi-level prevention system alone is not sufficient for RTI. A systematic, comprehensive, data-based decision making process is necessary to connect the pieces.

In a comprehensive RTI framework, data analysis occurs at all levels of RTI implementation, not just at the student level. For example,
- States may use RTI data to establish policy and guidance and allocate resources.
- Districts may use data to evaluate the effectiveness of RTI, establish policies and procedures, and allocate resources.
- Schools may use data to evaluate the effectiveness of their overall framework and the essential components, assess alignment among grade levels, and allocate resources.
- Grade-level teams may use data to evaluate core curriculum and instruction, identify students for secondary and tertiary instruction, and allocate resources.

Data analysis and decision making occur in all levels of prevention. For example, in primary prevention, the focus is the effectiveness of the core curriculum and instruction. With secondary and tertiary prevention, the focus is on student-level decisions, but also how well particular interventions work for the majority of students in the secondary and tertiary levels.

Districts and schools should have established routines and procedures, ideally in writing, for making decisions. Written procedures increase fidelity of the data-based decision-making process; ensure equity of resources among students, classes, and schools; and help train new teachers more efficiently. Teams should follow pre-established routines and procedures for making decisions. For example, data teams should meet at regularly scheduled intervals, such as monthly or bi-monthly, to systematically review data.

Districts and schools should also establish explicit decision rules for assessing student progress. This includes goal-setting procedures, changing instruction/interventions, referring students to special programs, and moving students to more or less intensive levels.

Schools can also use data to compare and contrast the adequacy of the core curriculum and the effectiveness of different instructional and behavioral strategies at all levels of prevention.
## Data-Based Decision Making: Types of Decisions

- Instruction
- Evaluate effectiveness
- Movement within the multi-level prevention system
- Disability identification (in accordance with state law)

These are the more common types of decisions that schools make.

1. **Instruction** – How effective is the instruction? What instructional changes need to be made?
2. **Evaluating effectiveness** – How do we know if the core curriculum, instruction for all students, interventions, and the RTI framework in general are effective?
3. **Movement within the multi-level prevention system** – How do we know when a student no longer needs secondary prevention or should move from secondary prevention to tertiary?
4. **Disability identification** – How do we know if the student should be referred and is eligible for disability identification? Decisions about disability identification should be made in accordance with your state’s laws.

Later in the presentation we will look in more detail at the types of data you might use within an evaluation to look at whether your RTI model is working. Remember that you are not only collecting different types of data, but are using data (screening and progress monitoring data) to inform your decision-making process within an RTI model.

Are there any questions about data-based decision making?

THINK-PAIR-SHARE

- How are you using data to make decisions within your RTI model?
  - Are there data teams?
  - Do you have explicit decision rules?
  - Are there routines and procedures in place for collecting and reviewing data?

Remind participants that they can use the Implementing the Essential Components Handout for note taking.

Think-Pair-Share

Think about how you are using data to make decisions within an RTI model.

Give participants approximately 20 seconds.

Pair and share with your neighbor/table and jot down your answer.

Give participants approximately 2–3 minutes.

Allow two or three pairs/tables to orally share their answers.
In review, as you can see in the graphic, the essential components of RTI are

1. **Screening**, or how we identify which students are at risk and need additional assessment and instruction
2. **Progress Monitoring**, or how we monitor whether students are responding to the instruction and supports we provide
3. **Multi-Level Prevention System**, or how we provide increasingly intense levels of supports to meet student needs
4. **Data-Based Decision Making**, or how we use data from those particular components to make decisions about student supports and program effectiveness

When evaluating your RTI model it is important to consider whether each of these essential components are in place and whether they are being implemented with fidelity. We will discuss fidelity more in the next section.
In this section we discuss fidelity of implementation and examples of tools that can be used to assess the implementation of RTI in schools and districts.

*Note: Much of this section is reinforcement from the RTI Implementer Series Multi-Level Module, which can be found at [http://www.rti4success.org/resourcetype/rti-implementer-series-module-3-multi-level-prevention-system](http://www.rti4success.org/resourcetype/rti-implementer-series-module-3-multi-level-prevention-system). If participants have already participated in the multi-level module you can use this section as a refresher.*

*Note: Fidelity and integrity are used interchangeably throughout this section.*
What Is Fidelity of Implementation?

This Ask the Expert video featuring Doug Fuchs addresses the following question: **We hear a lot about fidelity of implementation when talking about RTI. What does this really mean?** (5.17 minutes) Click play to watch the video.

Here is a sample from the video. Doug Fuchs: “Look, this is how we developed the program, this is the program. If you deliver the program the way we have detailed it, it’s a good bet that you will get results as we did. So what we’re really saying is, we’re encouraging fidelity of treatment implementation, meaning we’re encouraging you to implement our program the way we implemented it when we validated it. Importantly, this doesn’t mean that practitioners can’t take a validated instructional program, customize it to their own students and circumstances, and do better and have their children do even better than the children who participated in our research.”
Those in early implementation or interested in evaluating the impact of levels of implementation on changes in student outcomes also may be interested in evaluating the effectiveness of the level of implementation. It is important to remember that successful implementation does not guarantee changes in student success, so it will be important to measure that as well. Some indicators of successful implementation include the following:

• **Fidelity of implementation of RTI essential components.** If RTI is being implemented with fidelity, sites are implementing the essential components and the model with a high degree of fidelity. Later in this section we will discuss an Integrity Rubric developed by the National Center on Response to Intervention. This rubric allows schools to assess the integrity of implementation for each of the essential components as well as some overarching factors.

• **Effectiveness of the scale-up process, or the degree to which training, technical assistance, coaches, and so on were effectively implemented.** If the scale-up process is effective, sites will be implementing components with integrity and staff will have sufficient knowledge and skills to sustain implementation.

• **Impact of RTI on service delivery.** Changes may be observed in staff attitudes, scheduling, and delivery of instruction and supplemental supports, or in design of instructional supports.
Monitoring fidelity will allow us to understand whether the essential components have been implemented with integrity.

The best way to monitor fidelity is to measure it. Fidelity can be measured through self-report data, observations, or logs and lesson plans. We will discuss each of these in more detail.

Note: You may want to refer participants to the Measuring Fidelity Handout. During the following slides, participants can take notes about the different fidelity measures in the first columns of the handout. The remainder of the handout will be completed at the conclusion of the section.
Self-report data can include questionnaires, surveys, or interviews, and may provide an indicator of teacher knowledge as well as the context of implementation. Although these measures can be very efficient to conduct, they are often unreliable when used alone because they are subject to bias. For example, reports may include exaggerations or under-reporting in an attempt to make the responder look better.
Observation

- Types
  - Spot checks
  - Peer or administrator observations
  - Peer coaching
  - Item-by-item checklists of lesson components/rubrics

- Considerations
  - Develop checklists of critical implementation components
  - Record and listen to sessions at random
  - Least efficient but most reliable

Conducting observations can be done by developing checklists of critical implementation components, recording and listening to sessions at random, doing spot checks, conducting peer observations, and implementing peer coaching. Direct observations are the least efficient but most reliable form of fidelity measurement. Observations provide an outside and real-time perspective on what is going on rather than relying on the memory of person implementing RTI. The reliability of observations can be increased by ensuring the observers are trained, using multiple observers with a goal of achieving high inter-rater reliability (or high level of similar responses), and using detailed checklists and observation tools that anchor responses in specific behaviors or practices. While conducting observations eliminates the potential bias caused by the individual reporting on him- or herself, it may create situations where the individual being observed acts differently when being observed.
Reviewing logs/lesson plans and student work allows evaluation of what was done. It could include looking at the content covered and student progress. Reviewing logs, lesson plans, and student work is moderately efficient and moderately reliable. It provides less information about delivery, dosage, and adherence to scripts (if applicable) than other measures of fidelity, however.
Have participants turn to the Measuring Fidelity Handout and complete the handout with their teams. The information in the first three columns is a review of the information just covered in the slides, and the last two columns ask teams to think about how they are currently measuring fidelity and what they might do in the future. Provide 5–10 minutes for this activity. You can also suggest that participants take notes in the first few columns as you go through information in the slides.

Instructions: Use this handout to track information on fidelity and work with your team to determine your current practices to measure fidelity of implementation and make plans for what other measures you could implement to measure fidelity. Instruct teams to select an essential component (screening, progress monitoring, multi-level prevention system, or data-based decision making) to use as a lens for this activity.

Remind teams after group time that they can continue this process for the remaining essential components.
We just discussed three main ways to assess fidelity of implementation: self-report; observation; and logs, lesson plans, and student work. We are now going to look at some examples of what states use to evaluate implementation of RTI components and the RTI Process.

### Examples of Measures and Tools for Evaluating Implementation of RTI Components and the RTI Process

- Fidelity-of-implementation forms
- Self-assessment rubrics
- Interview protocols
- Product review forms
- Free RTI surveys and checklists
- Direct observation instructions or rubrics
The National Center on Response to Intervention developed the integrity rubric and accompanying worksheet as a means for schools and districts to assess or self-assess their progress in implementing RTI.

**Summarize the slide.**

NCRTI Integrity Rubric

- Intended for use by individuals responsible for monitoring school-level fidelity of RTI implementation or as a measure of self-assessment.
- Provides descriptions of three levels of potential ratings (1, 3, or 5) across each factor.

The RTI Essential Components Integrity Rubric and the RTI Essential Components Integrity Worksheet are for use by individuals responsible for monitoring the school-level fidelity of Response to Intervention (RTI) implementation. They may also be used by schools for self-appraisal; however, they were not designed for compliance monitoring and therefore should not be used for this purpose. The rubric and the worksheet are designed to be used together, and are aligned with the essential components of RTI.

The sections of the rubric are exactly aligned with the sections of the worksheet.

The rubric provides a five-point rating scale and descriptions of practices that would score a 1, 3, or 5. If a school’s practice seems to fall between the described ratings, the school is assigned a rating of 2 or 4. For example, a school judged to be performing at a level higher than the rubric describes for a 3 rating but not quite at the level described for a 5 would receive a 4 rating.

Let’s take a look at the rubric to illustrate this.

Note: As we mentioned earlier there is a difference in the reliability of self-report data and external observation. Using an external evaluator to observe and fill out the rubric is more objective and reliable than relying on self-report data provided in the worksheet. Filling out the worksheet still provides benefits for the school as they are able to see what evidence is available to support their assessment of implementation.
Here is the first page of the integrity rubric.
As you can see, the rubric describes a rating of 1 as **insufficient evidence that the screening tools are reliable; or that correlations between the instruments and valued outcomes are strong; or that predictions of risk status are accurate.**
Review the slide.

NCRTI Integrity Framework Worksheet

- Intended for use by RTI coordinators or evaluators with extensive RTI experience
- Provides space to develop a narrative rationale for each rating
- Data collected through interviews or site visits
Here is an example of the worksheet.
As you can see, the information for screening tools includes asking the following questions:

1. **What tools do you use for universal screening?**

2. **When your school selected the screening tool(s), how much attention was paid to the evidence from the vendor on the validity, reliability, and accuracy of the tool?**

3. **Does your school have documentation from the vendor that these tools have been shown to be valid, reliable, and accurate (including with subgroups)?**

4. **Do you have reason to believe that the screening tool(s) that you use may have issues with validity, reliability, or accuracy (including with subgroups)? If so, please explain.**
The RTI State Database provides resources on a number of topics related to Response to Intervention (RTI), including measures of fidelity. Resources were compiled by the National Center on Response to Intervention (NCRTI) in an effort to share examples and information across states. They were gathered from public sources (e.g., websites, SPPs, APRs) and are intended only to provide examples—not recommendations—of RTI implementation and fidelity measures in the field.

In the next few slides, we will provide examples from a number of states. Each of these tools can be found in the state database. There are many more resources available online.

Note: The RTI State Database was last updated in spring 2012. More recent information from states may be found on their state website. The trainer should determine whether participants would like to have any of the following examples as standalone handouts.
This is an example of Colorado’s RTI Implementation Rubric for the district level. The RTI Implementation Rubrics are a set of rubrics that serve as an overview of implementation for RTI.

Rubrics are available for the classroom, school, and district level. Each rubric describes what RTI looks like across the six components of RTI (problem solving, curriculum & instruction, assessment, leadership, family & community partnering, positive school climate) and across four growth stages (emerging, developing, operationalizing, optimizing).

The purpose of the rubrics is to

1. Serve as an informational resource (i.e., blueprint, roadmap of RTI implementation)
2. Measure fidelity of RTI implementation
3. Assist with planning for an action plan or school improvement plan

Visit http://www.cde.state.co.us/rti/ to find each of the rubrics focused on the district, school, and classroom level as well as additional materials. The state began developing these rubrics in 2010, and they have been available since winter 2011.
This is an example of a self-report form that was developed by the state of Connecticut.

For each item in the self-assessment, an individual

- Records the rank priority—high, medium, or low
- Provides evidence to support each item
- Records the rank level of implementation—not yet, initial, partial, or full

This is an example of an interview form developed by Connecticut to help evaluate the RTI process. An interviewer may ask similar questions to staff in different roles, such as the administrative team, teachers, and support staff, to obtain multiple perspectives on the RTI model in the district or school.

Interviews also should be conducted with the district and school data team. Questions may include the following (click for animation):

- What is the purpose of your team?
- What are the activities of the team?
- What is the expectation of the team?
- How often do you meet? How long is the meeting?
- How is the agenda for the meeting determined?
- How are decisions determined?
- How are strategies for student improvement determined? How are they evaluated?
- How does the data team influence classroom/school-wide practice (e.g., coaching teachers, support personnel)?
- Give an example of how the data team supports improvement in student outcomes (e.g., academic or behavioral).

This resource can be found here: http://www.sde.ct.gov/sde/lib/sde/word_docs/cali/srbi_observation_interview_protocol_october_2008.doc. This document was developed in 2008.
This is an example of a permanent product form developed by Connecticut. This form is used to examine screening and progress monitoring data. The assessments that are used within the district are listed and information is provided about each assessment.

You can view this resource here: http://www.sde.ct.gov/sde/lib/sde/word_docs/curriculum/cali/assessment_inventory.doc. It was last revised in February 2012.
The next two slides provide two examples of methods for conducting direct observations, as developed by Connecticut.

The first example lists key areas that the observer should be looking for on his or her walkthrough.

- **The observer would write down evidence of practice, notes, or comments for each these key areas.**
- **The observer also may record the readiness level (beginning, developing, proficient, or exemplary) observed for each of these key areas.**

Additional information on observations can be found here: http://www.sde.ct.gov/sde/lib/sde/word_docs/cali/srbi_observation_interview_protocol_october_2008.doc. This observation and interview protocol was developed in October 2008.
Example—Connecticut’s Evaluation of Districts’ RTI Procedures: Data Team Observation

In this example, observers would check the box if they saw evidence of the key area during the observation period.

<table>
<thead>
<tr>
<th>Data Team Observation (30 minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations:</td>
</tr>
<tr>
<td>- Check all that apply:</td>
</tr>
<tr>
<td>- Prepared agenda</td>
</tr>
<tr>
<td>- Consistent schedule (weekly/bi-weekly)</td>
</tr>
<tr>
<td>- Facilitator</td>
</tr>
<tr>
<td>- Record or minutes taken</td>
</tr>
<tr>
<td>- Focus on student work/outcomes</td>
</tr>
<tr>
<td>- Collecting or analyzing data on all students</td>
</tr>
<tr>
<td>- Setting SMART goals</td>
</tr>
<tr>
<td>- Selecting effective teaching strategies</td>
</tr>
<tr>
<td>- Determining effectiveness of strategies</td>
</tr>
<tr>
<td>- Determining specific actions for subsequent data team meeting</td>
</tr>
</tbody>
</table>

The second example is an observation checklist that an observer would use during an observation of the data team meeting process.

- The **observer would simply check the box if he or she saw evidence of the key area during the observation period**.

Additional information on observations can be found here: http://www.sde.ct.gov/sde/lib/sde/word_docs/cali/srbi_observation_interview_protocol_october_2008.doc. This observation and interview protocol was developed in October 2008.
This Delaware form is used for evaluating fidelity of implementation of core instruction.

- Three levels of fidelity of implementation—A, B, and C
- Potential team action steps to improve fidelity of implementation

You can find this resource here: http://www.doe.k12.de.us/infosuites/staff/profdev/rti_files/District%20RTI%20Planning%20Guide.doc. This resource was developed in July 2007.
CONSIDERATIONS FOR DEVELOPING AN EVALUATION PLAN

In this section, we will discuss some important things to consider when developing an evaluation plan.
Some of the key considerations are—

Read the slide.
Data should be collected to answer your evaluation questions. The following slide gives some examples of what you may collect data on.

*Review contents of the slide.*

There are a number of data sources that can be used to collect these types of data—for example, screening data, progress monitoring data, high school GPA, suspension and referrals, and so on.
### Data Collection: Key Questions

- **What data do we need to collect?**
- **Can we use existing data? If so, what sources?**
- **Do we need to collect additional data? If so, how will we collect the data?**
- **Can we use similar data sources across all sites?**

---

*Review contents of the slide.*

- **What data do we need to collect?** You will need to consider what type of data you will use to answer your evaluation questions and whom you will collect the data from. Consider whether you will collect data from all schools or just some of the schools in the district.
- **Can we use existing data? If so, what sources?**
- **Do we need to collect additional data? If so, how will we collect the data?** Schools collect a lot of data for a lot of different school initiatives. It is important to limit the burden for schools as much as possible.
- **Can we use similar data sources across all sites?** If possible, it is best to use similar data sources across all sites. This will make analysis easier and will allow you to aggregate data across sites.
Data Collection: Methods

- Consistency across schools’ data reporting systems and data collected
  - Use a common language across all schools
  - Collect consistent data across schools
  - Use consistent cut scores and benchmarks
- Multi-method/multi-measure assessment is preferred

To evaluate RTI across schools in a district, it is preferable that there be **consistency across schools’ data reporting systems and data collected** at the various schools. It is important to **use common language, collect consistent data across schools, and use consistent cut scores and benchmarks**. To minimize problems, improve data reporting procedures, and increase efficiency, it is important for districts to consider a common, integrated data system. If schools do not use consistent measures, it is difficult to compare data from one school to another school. For example, if one school has a cut score that is much lower than other schools’, it may appear that more of their students are meeting the cut score, but in fact their students may not be performing better than the other schools. See the RTI Implementer Series: Screening Module at www.rti4success.org for slides explaining this.

Evaluations should draw data from multiple sources that relate to your outcomes of interest. A single-source assessment will most likely not present a full and complete picture of RTI implementation and effectiveness.
Data Collection: Methods

<table>
<thead>
<tr>
<th>Evaluation Question</th>
<th>Success Indicators</th>
<th>Type of Data Needed</th>
<th>Submission or Collection Methods</th>
<th>Timeframe of Data Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is it we want to know?</td>
<td>What is an indicator of what we want to know?</td>
<td>What type of data will provide information for the success indicator?</td>
<td>How will the data be collected and analyzed?</td>
<td>What is the timeframe for data collection?</td>
</tr>
</tbody>
</table>

A data collection matrix can be a useful part of developing an evaluation plan. Some of the information you may consider recording includes the following:

**Evaluation questions. What is it that we want to know?** For example, how effective is RTI in reducing the number of at-risk students?

**Success indicators. What is an indicator of what we want to know?** For our previous example, we may have identified risk level as a potential success indicator to understand if RTI has helped reduce the number of at-risk students.

**Type of data needed. What type of data will provide information for the success indicator?** For example, We will need the number of students identified as at risk using screening data at each benchmark period.

**Submission or Collection Methods. How will the data be collected and analyzed?** The data will be submitted and collected through the published data system that houses the screening and progress monitoring data.

**Timeframe of Data Collection. What is the timeframe for data collection?** The timeframe for the data collection will be the fall, winter, and spring screening periods. A place is also provided to highlight potential concerns or challenges that may be faced with this data collection method.
Staffing and Expertise

- Who will oversee the evaluation process?
- Who will be responsible for collecting and analyzing data?
- Do we have sufficient internal staff expertise to design and conduct an evaluation?
- Do we need external expertise? What is available?

*Review contents of the slide.*

While the individuals involved in the evaluation process may vary from school to school, it may be good for teams to consider including people with expertise in interpreting assessment data. This may be school psychologists or assessment directors with a background in statistics or educational psychology.
Funding

- What funding resources are available to support the ongoing evaluation of RTI?
- What funding is available to support knowledge development or engage experts?
- What funding is available to address any necessary changes to the RTI process?

Review contents of the slide.
Evaluation Tools

- What tools are currently available?
- Can we use existing tools? Do we need to adapt or develop new tools?
- What resources do we have for analyzing and reporting the results?
- Do we have sufficient training on how to use the tools and reporting information we select?

Note: This will be covered further in later sections of the presentation.

Review contents of the slide.
Timeframe

- How much time is needed to plan for evaluation?
- When will data collection begin?
- Will multiple years or a single year of data be used for analysis?
- When will data analysis occur (single time point or multiple time points) and when will the results be reviewed?

Review contents of the slide.
Data Analysis

- How will we aggregate data (across students, classes, grades, schools)?
- Do we have consistent benchmarks (across grades, within the district)?
- How many years of data will we consider?
- Will we use a data system to support analysis? If so, which one?

There are many different ways to analyze evaluation data. This slide provides some questions that you may consider when planning for data analysis. A later section will show examples of different types of analysis and what they can tell you.

*Review contents of the slide.*

*Note: More information on data analysis will be presented in a later section of the presentation.*
Sharing of Results

- How will we report the results?

- How will we disseminate the results and next steps?
  - Parents
  - School board
  - Teachers
  - Administration
  - School staff
  - Others

Review contents of the slide.
Think-Pair-Share

Think about how you will share the results of your evaluation, or how you currently share your evaluation data.

*Give participants approximately 20 seconds.*

Pair and share with your neighbor/table and jot down your answer.

*Give participants approximately 2–3 minutes.*

Allow two or three pairs/tables to orally share their answers.
Personal Journaling Opportunity:

Provide participants about 5 minutes to individually journal thoughts about the questions posed on the slide.

Reflection: Guiding Questions for Evaluation

- What do we want to know?
- What data do we have that we can use to evaluate (e.g., fidelity data, student outcome data)?
- Do we need to collect additional data? If so, what type?
In thinking about student outcomes, there are a number of areas to consider. The success indicators that we focus on will depend on the type of information we are looking for. Success indicators dictate the type of data collection and analysis that should be used in the evaluation process. For example, if the purpose is to demonstrate the effectiveness of RTI in improving student outcomes, then performance on an outcome measure may be an appropriate success indicator. To show the impact of RTI in reducing the number of students identified as at risk, a success indicator may look at the students identified by the screening measure as at risk or not at risk across the year or multiple years. It also might consider the percentage of students moving between levels of instruction, and the direction of the movement.

We are going to discuss a number of different success indicators in this section. After each indicator, there will be time for teams to work together to discuss some key questions related to that indicator.

*Have participants take out the Evaluation Plan Graphic Organizer for this section.*
As we think about the different ways that we can assess indicators of success, we also need to consider some overarching ideas about data analysis. These first two slides will focus on the different ways that we might analyze the data in our evaluation to look at different indicators of success and what systems we can use to make analysis easier.

One consideration is whether we are looking at student-, grade-, school-, or district-level outcomes. Aggregating data across schools, grades, years, and so forth allows us to look at the overall picture of the impact of RTI. Data at the individual school and district level are subject to variability in fidelity of implementation and outcomes. By aggregating data across a larger sample, we can smooth out outliers. Because we are aggregating data, it is important to pay attention to both the mean (average) and the range (spread) of scores.
We also need to consider whether we will be looking at the changes within a given year (fall, winter, spring) or across multiple years. When we look at data within a given year, we are comparing the same students in the given grade level or school across the year. In looking across multiple years of data, we are considering different cohorts of students. Looking at data across multiple years can help us to see whether there are differences in outcomes resulting from the number of years that RTI has been implemented in the school. This is especially important for evaluating the whole RTI model.

Another analysis decision we need to consider is whether we are just looking at an individual school or at multiple schools that are implementing RTI, or are comparing outcomes for school implementing RTI with schools that are not implementing RTI.
To evaluate the effectiveness of your RTI model, it may be important to use a data system. Using electronic data in preference to written data can help to increase efficiency. Published data systems and Excel spreadsheets can be used for this purpose, but can have different advantages and disadvantages.

Review contents of the slide.
All measures of success are related to data-based decision making, but two common indicators of interest are

- Student outcomes (e.g., scores on end-of-year assessments, number of office referrals)
- Patterns of special education referral

Review the slide.
Student Outcomes

- Student outcomes can include
  - State assessments
  - Discipline/behavioral referrals
  - Graduation rates
  - Retention rates
  - Dropout rates
- The percentage of students achieving, nearly meeting, or exceeding standards on outcome measures should increase across years.

States, districts, and schools are often very interested in understanding whether RTI is helping to improve student proficiency on state assessments. If RTI is working, the percentage of students meeting or exceeding standards on measures such as the state assessment should increase across years, across cohorts of students, and across years within cohorts. But using this indicator alone makes it difficult to determine whether RTI implementation is the contributing factor.

Examples of other outcome measures of interest are discipline/behavioral referrals, graduation rates, dropout rates, and retention rates.

How often the outcome measure of interest is assessed may depend on the type of measure. For example, state assessments, graduation rates, dropout rates, and retention rates may be measured annually, and discipline or behavior referrals may be assessed more regularly.
We are now going to look at two examples of ways to look at changes in student achievement on state assessments. The first example looks at changes in student proficiency across three years of RTI implementation and across three grade levels (i.e., cohorts). The second example compares changes in proficiency for schools implementing RTI with schools not implementing RTI.
This graph depicts example data for student proficiency on the state assessment across three years across all RTI schools (examples are included for three different grades). Specifically, this graph shows the percentage of students who met or did not meet proficiency on the assessment for students in third grade, fourth grade, and fifth grade during the first, second, and third year of RTI implementation. We know that full implementation of RTI may take 2–4 years and that each year schools innovate and refine their models to meet the needs of their students. An effective model shows improvement, with more students reaching proficiency over time.
Think-Pair-Share

Think about what the graph tells you about student achievement in the following grades. What trends do you see for students in third grade? What about fifth grade?

Give participants approximately 20 seconds.

Pair and share with your neighbor/table and jot down your answer.

Give participants approximately 2–3 minutes.

Allow two or three pairs/tables to orally share their answers.

Note: All data in this presentation are fictitious.
In this example, we are looking at a comparison of schools in the district that are implementing RTI with those that are not implementing RTI. The example shows data for the same schools across three years. In this example, the RTI schools began implementing RTI in the 2008–09 school year. As we mentioned previously, we know that full implementation of RTI takes 2–4 years, so not only are we looking to see the proficiency rates for students at the schools implementing RTI compared with students at schools not implementing RTI, but also how the percentage of students reaching proficiency has changed during RTI implementation.

Note: All data in this presentation are fictitious.
Think about what the graph tells you about student proficiency at schools implementing RTI in comparison to those not implementing RTI.

Give participants approximately 20 seconds.

Pair and share with your neighbor/table and jot down your answer.

Give participants approximately 2–3 minutes.

Allow two or three pairs/tables to orally share their answers.

Note: All data in this presentation are fictitious.
### Evaluation Plan Graphic Organizer

Work with your team to fill out section 1 of the graphic organizer, focusing on these questions:

- What are you currently doing to evaluate student outcome measures?
- What processes can you implement to evaluate student outcome measures?

---

*Have participants turn to the Evaluation Plan Graphic Organizer and complete section 1, focused on student outcome measures, with their teams. Provide 5–10 minutes for this activity.*

*Instructions: The handout includes sections focused on the considerations that we discussed in the prior part of the presentation. You will see questions related to data collection, staffing and expertise, funding, evaluation tools, timeframe, data analysis, and data sharing. With your team, answer the questions on the basis of what you are doing to evaluate student outcome measures and what your plans are for the future.*
Although it may not be the primary focus of RTI for some schools, another area of interest is how patterns of special education referrals change with RTI implementation. Research suggests that we should see fewer inappropriate referrals to special education. This would be logical, because students who are responding to the interventions that are being presented, whether at a secondary or tertiary level, may be less likely to be referred for special education. As a result, we would expect that special education evaluation would become more efficient. Students who are referred for evaluation out of an RTI process should have a much higher probability of being found eligible then those referred outside the context of an RTI process. This is a result of controlling for inadequate instruction and identifying risk earlier to facilitate prevention efforts.

Disability Identification

To ensure that underachievement in a child suspected of having a specific learning disability is not due to a lack of appropriate instruction in reading or math, the group must consider, as part of the evaluation, what is described in 34 CFR 300.304 through 300.306:

- Data demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in regular education settings, delivered by qualified personnel.
- Data-based documentation of repeated assessments of achievement at reasonable intervals reflect formal assessment of student progress during instruction, which was provided to the child’s parents.

Remember that federal law states that “to ensure that underachievement in a child suspected of having a specific learning disability is not due to lack of appropriate instruction in reading or math, the group must consider” two things. The first is whether “data demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in regular education settings, delivered by qualified personnel.” Screening data that look at the growth rate of all students can help us answer this question. Progress monitoring data that can be shared with parents can support the next piece: “data-based documentation of repeated assessments of achievement at reasonable intervals.”

Authority: The requirements for using a process based on a child’s response to scientific, research-based intervention when determining whether the child has a specific learning disability are found in the regulations at 34 CFR §§300.307, 300.309, and 300.311.
State Regulations for Disability Identification

- States have different regulations for referrals for disability identification.
- Contact the State Department of Education for additional information.

Review slide.

Note: Different states have different regulations for disability identification. This slide provides the presenter an opportunity to include state-specific information about the referral process in the state of interest. If this is a cross-state training, you may want to delete this slide or provide examples from a couple of states.
Have participants turn to the Evaluation Plan Graphic Organizer and complete section 2, focused on patterns of special education referrals, with their teams. Provide 5–10 minutes for this activity.

Instructions: The handout includes sections focused on the considerations that we discussed in the previous section of the presentation. You will see questions related to data collection, staffing and expertise, funding, evaluation tools, timeframe, data analysis, and data sharing. With your team, answer the questions on the basis of what you are doing to evaluate patterns of special education referrals and what your plans are for the future.
Using screening data, we can look at the changes in the number of students who are identified as at risk or not at risk based on screening benchmarks.
Universal screening measures are used in an RTI model to identify students who are at or below a benchmark. A benchmark can be determined in different ways by using different kinds of metrics. When we talk about students who are at risk, we talk about students who have not met the benchmark and are identified as at risk and those who have met the benchmark and therefore identified as not at risk. Over time, we should see an increase in the number of students who achieve the benchmark levels over the course of a year and we should see an associated decrease in the number of students who are at risk. The goal for most RTI models is to have about 80% of their students reaching the benchmark or cut score for the screening measure. If fewer than 80% are reaching the benchmark, emphasis should be placed on improving the core or general education curriculum and instruction.

<table>
<thead>
<tr>
<th>Indicators of Success: Using Screening Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ Percentage of students achieving the benchmark should increase over the year.</td>
</tr>
<tr>
<td>• Example: Oral Reading Fluency benchmarks (AIMSweb, DIBELS, local norms)</td>
</tr>
<tr>
<td>▪ Percentage of students achieving benchmark across years should increase toward ceiling levels.</td>
</tr>
<tr>
<td>• The goal is to have at least 80% of students identified as not at risk</td>
</tr>
</tbody>
</table>
The “about 80%” that is often cited as a goal for schools is based on the public health model. Although it provides a goal, it may be unrealistic for some schools.

- For example, this goal may be challenging for schools with highly transient populations with a large percentage of students coming and going. These schools have students who suddenly emerge as at risk who were never exposed to core curriculum and interventions that may have been provided in kindergarten, first grade, or an early intervention period. Therefore these students may not have the same abilities and skills as other students. This of course would increase the risk level.

- Other examples are schools with a high percentage of students living in poverty, or that have a high percentage of English language learners.

These schools should still aim for 80%, but should look at their baseline rates as a starting point in order to set reasonable goals and expectations.
Example: Establishing Goals for Benchmark at School A

- A small urban community, high poverty, failing school
- When an RTI model was introduced into that school
  - First universal screening revealed that 16% of students in the early grades met benchmark: 60–70% did not.
  - Rather than establish an unrealistic expectation that at least 80% of students reach benchmark over the course of several years, they established a goal of 50% at benchmark in year 1, and 60% at benchmark in year 2.
  - While the school did not reach 80% at benchmark, it improved over a three-year period to about 55% at low risk and only 20% at risk.

For example, School A is located in a small urban community, with a high percentage of students living in poverty. It had been recognized as a failing school for multiple years and had many students not achieving benchmark on state tests. When an RTI model was introduced into that school, results from the first universal screening revealed that there was a baseline rate of students at benchmark (i.e., not at risk) of 16% in the early grades (with a range of 10 to 20% at benchmark across the grades), with 60–70% of students at risk. To set an expectation of getting 80% of students to benchmark over the course of several years would be unrealistic in this setting. Establishing a goal of 50% at benchmark in year 1, and 60% at benchmark in year 2, however, would be more reasonable, and indeed, this school was successful in meeting that goal. The school was unable to reach 80%, but the risk level improved over a three-year period to only 20% at risk.
We are now going to look at three examples of how to look at students achieving the benchmark through screening data. In aggregating data across these examples, it is important that a consistent cut score be used across schools. The first example looks at changes in risk level (both mean and range) within the same school year, the second example compares different years of RTI implementation, and the third example compares changes in the percentage of students identified as at risk or not at risk at schools implementing RTI versus schools not implementing RTI.

Remember, looking across years is important for understanding how the RTI model is working in the school.
In the first example, we are going to look at changes within the same year aggregated across seven schools in a district. Remember that aggregating data across the schools in the district allows us to understand the effectiveness of the RTI in general. Further analyses can help determine the effectiveness by individual school. When looking at the risk status, it is important to take into account the mean and the range. First we look at a graph that shows average percentage of students meeting and not meeting the benchmark by grade level (first and second grade) across seven elementary schools within one year. The following slides will show examples of the range for the percentage of students identified as at risk and not at risk.

The graph shows the percentage of students who were identified by the screener as at risk or not at risk. As you can see, across the year, the percentage of students identified as not at risk increased from 73% in the fall to 84% in the spring according to the word identification fluency (WIF) screener. This suggests that fewer students were being identified as in need of supplemental supports.

*Note: All data in this presentation are fictitious.*
Think-Pair-Share

Think about what the graph tells you about second graders’ progress in these elementary schools and what questions may you raise based on the data.

Give participants approximately 20 seconds.

Pair and share with your neighbor/table and jot down your answer.

Give participants approximately 2–3 minutes.

Allow two or three pairs/tables to orally share their answers.

With second-grade oral reading fluency (ORF), you can see that there is no change from fall to winter and that there is a slight decline between the winter and spring. Overall, these two graphs illustrate a pattern of change in risk level on one measure for first-grade students but not for second graders.

Note: All data in this presentation are fictitious.
This graph shows an example of how you could look at the range. Remember, when you are aggregating data, it is important to look at the range, because some schools in the district might be outliers. The white boxes represent the range of the percentage of students identified as not at risk by the screener for five schools during the fall, winter, and spring. The pink line represents the change in the top of the range—the increase or decrease in the highest percentage of students not identified as at risk; the blue line indicates the change in the bottom of the range—the increase or decrease in the lowest percentage of students not identified as at risk. For example, in the fall, the percentage of students identified as not at risk on the word identification fluency screener in first grade ranged from 29% to 85%, indicating large differences among schools in the district. The pink line shows that despite a decrease from the fall to the winter across the year, from fall to spring there was an increase in the top of the range—an increase in the highest percentage of students not identified as at risk by the Word Identification Fluency screener (pink and blue lines).

*Note: All data in this presentation are fictitious.*
Here is a different way to look at the risk status. In the previous graphic, we looked at the percentage of students identified as not at risk. In this graph, we are looking at the range of the percentage of students identified as at risk by the screener for the same five schools.

Note: All data in this presentation are fictitious.
In this example, we are looking at a graph that shows the risk status of students in the same grade across three years of RTI implementation. In example 1, we were looking at the same students across the entire year. In this example, we are using three years of data for each grade and therefore are looking at three different cohorts of students.

This graph shows results for students in first grade and second grade during the first, second, and third years of RTI implementation. We know that full implementation of an RTI may take 2–4 years and that each year schools innovate and refine their models to meet the needs of their students. In an effective implementation, the percentage of students identified as not at risk should increase while the percentage of students identified as at risk should decrease over time. In this graph, this pattern is shown for both first and second grade.

Note: All data in this presentation are fictitious.
In this example, we are looking at a comparison of schools in the district that are implementing and not implementing RTI. The previous example showed data for the same schools across three years. In this example, the RTI schools began implementing RTI in the 2008–09 school year. As in Example 2, we are able to look at risk-level changes across RTI implementation (i.e., across cohorts of students).

The data presented here indicate that over the course of three years, the percentage of students not identified as at risk increased for the first graders in schools implementing RTI, but for schools not implementing RTI, the percentage of students not identified as at risk varied across the years, with some improvement shown from the first to the second year but a decline from the second to the third year.

Group Question: **What does the graph say about the change in risk level across years of RTI implementation for schools implementing RTI?**

Answer: The graph shows that a lower percentage of students is being identified as at risk across the three years of RTI implementation for those schools implementing RTI.

Some schools and districts may use multiple screening measures for the same grade. It is important to consider all of these measures when conducting evaluations.

*Note: All data in this presentation are fictitious.*
Have participants turn to the **Evaluation Plan Graphic Organizer** and complete section 3, focused on benchmark risk status with screening data, with their teams. Provide 5–10 minutes for this activity.

Instructions: The handout includes sections focused on the considerations that we discussed in the previous section of the presentation. You will see questions related to data collection, staffing and expertise, funding, evaluation tools, timeframe, data analysis, and data sharing. With your team, answer the questions based on what you are doing to evaluate changes in student risk levels using screening data and what your plans are for the future.
In the progress monitoring module we discussed decision-making processes for determining whether students were making adequate progress. It is important to consider this information when thinking about whether students are closing the gap or whether they need additional supports in order to reach targeted outcomes.
In order to understand the concept of movement within levels of intensity, it is important to understand the term “rate of improvement” (ROI). ROI is the rate of growth made by a student between one point in time and another. There are three key terms to understand when using ROI:

- **Typical ROI**—average ROI for the year for a typically performing student at that grade level. This information is typically the national norm provided by progress monitoring tools.
- **Targeted ROI**—the ROI the student needs to attain in order to close the gap or meet the benchmark. This is the goal that is set for the student to achieve the benchmark.
- **Attained ROI**—the ROI the student obtained over the course of the instruction period. This is the student’s actual performance.

Remember, you can aggregate the data across grade levels and schools just as we have done with our data for risk levels and movement between levels.
This table provides an example of the national norms for typically developing students at the grade level at which the student is being monitored. The table provides the average rate of weekly increase from a national norm chart. This national norm also can be thought of as the typical ROI.

These are sample national norms for weekly improvement based on one assessment tool. Remember, the national norms for weekly improvement you select will depend on which tool you have selected, as well as other factors, such as additional analysis conducted for your specific population. Teams should check with their assessment vendor for more accurate national norms.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Reading—Slope</th>
<th>Computation CBM—Slope for Days Correct</th>
<th>Concepts and Applications CBM—Slope for Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.8 (WIF)</td>
<td>.35</td>
<td>No data available</td>
</tr>
<tr>
<td>2</td>
<td>1.5 (PRF)</td>
<td>.30</td>
<td>.40</td>
</tr>
<tr>
<td>3</td>
<td>1.0 (PRF)</td>
<td>.30</td>
<td>.60</td>
</tr>
<tr>
<td>4</td>
<td>.40 (Maze)</td>
<td>.70</td>
<td>.70</td>
</tr>
<tr>
<td>5</td>
<td>.40 (Maze)</td>
<td>.70</td>
<td>.70</td>
</tr>
<tr>
<td>6</td>
<td>.40 (Maze)</td>
<td>.40</td>
<td>.70</td>
</tr>
</tbody>
</table>

Note: These figures may change with further RTI research and are specific to one assessment tool.
We will now look at two examples of how you might evaluate movement within levels of intensity. This information is typically reflected through progress monitoring data. The first example looks at a single student’s typical, targeted, and attained ROI for both a student who is not closing the gap between targeted and attained and one who is closing the gap. The second example uses aggregated data across grade levels in a single school to compare targeted and attained ROI.

Ways to Look at Movement Within Levels of Prevention

1. Comparisons of typical, targeted, and attained ROI for a single student
   a) Not closing the gap
   b) Closing the gap
2. Aggregated data comparing targeted and attained ROI by grade level for students in a single school
On this graph, the trend line is flatter than the goal line. A trend line below the goal line indicates that student progress is inadequate to reach the end-of-year performance goal or benchmark. The instructional program should be tailored to bring this student’s scores up. Remember, the goal should never be decreased.

*Note: The typical ROI should be provided by the tool publisher. The language used may vary. In some instances, the targeted ROI and typical ROI may be the same. If the goal has been raised because the typical ROI does not appear to be challenging enough, the targeted ROI might be higher than the typical ROI.*

*Note: This graph shows a certain instructional period representing 14 weeks of the school year. For this example, the end-of-year benchmark will occur at the 14th week. In reality, the weeks of instruction would represent the total number of weeks between when progress monitoring begins and the end of the instructional period that you are setting a goal for. This may vary by school and student.*
Example 1b—Student Closing the Gap (Actual and Targeted ROI)

On this graph, the trend line or actual ROI is steeper than the goal line or targeted ROI. The student is showing increasing scores; therefore, the student’s end-of-year goal needs to be adjusted to provide more of a challenge. The student also could be moved to less intensive support if he or she meets the grade-level benchmark.

Note: This graph shows a certain instructional period representing 14 weeks of the school year. For this example, the end-of-year benchmark will occur at the 14th week. In reality, the weeks of instruction would represent the total number of weeks between when progress monitoring begins and the end of the instructional period that you are setting a goal for. This may vary by school and student.
Now that you understand targeted and attained ROI, we can use these data to evaluate the impact of RTI on student performance. This summary graph shows aggregated data across Grades 2 and 3 for targeted versus attained levels for students in secondary prevention in one RTI school. Note that the targeted ROI (the goal) and the typical ROI are the same for Grade 2, but the targeted ROI is higher than the typical ROI or national norm in Grade 3.

Let’s take a closer look. In third grade, there were 18 students receiving secondary instruction. The typical ROI, shown by the red line, for students in third grade is 1.0 wcpm per week. Across the 18 students receiving secondary instruction, the average targeted ROI (green bar) was 1.25 wcpm per week. Together, these students gained an average attained ROI (purple bar) of 1.66 wcpm per week. This group of students closed the gap, indicating that the secondary model for third grade is effective. The attained ROI (1.66 words correct per minute per week) exceeded the typical ROI (1.00 words correct per minute per week) and the targeted ROI (1.25 words correct per minute per week).

Here is another example. In second grade, there were nine students receiving secondary instruction. The typical ROI (red line) for students in second grade is 1.5 wcpm per week. Across the nine students receiving secondary instruction, the average targeted ROI (green bar) was 1.5 wcpm per week. Together, these students gained an average attained ROI of 1.14 wcpm per week. This group of students did NOT close the gap, indicating that the secondary model for second grade is not effective. Their attained ROI (1.14 wcpm per week) was below both the typical ROI (1.5 wcpm per week) and their targeted ROI (1.5 wcpm per week).

Note: You will notice that the typical rate of improvement line decreases across grade levels. This decrease in rate of improvement is a result of growth rates slowing down over years. In other words, it is easier to make more progress in lower grades than in higher grades.
Think about the data for first-grade students. The typical ROI for first-grade students is 1.8.

- How are the students in first grade doing?
- What does this say about the secondary model for first grade?

Give participants approximately 20 seconds.

Pair and share with your neighbor or table and jot down your answer.

Give participants approximately 2–3 minutes.

Allow two or three pairs or tables to orally share their answers.
Have participants turn to the Evaluation Plan Graphic Organizer and complete section 4, focused on movement within levels using progress monitoring data, with their teams. Provide 5–10 minutes for this activity.

Instructions: The handout includes sections focused on the considerations that we discussed in the previous section of the presentation. You will see questions related to data collection, staffing and expertise, funding, evaluation tools, timeframe, data analysis, and data sharing. With your team, answer the questions based on what you are doing to evaluate student progress within prevention levels and what your plans are for the future.
Not only can we look at movement within prevention levels, as seen in the previous section, but we also can look at the movement between levels of prevention.

*Review slide.*
In an effective model, we want our data to reflect the majority of growth in one direction, the majority of students moving to less intensive instruction. With this analysis, we have to assume that the levels of intensity across implementation sites are similar. In other words, what school A refers to as secondary prevention is similar to school B’s secondary prevention level.

Keep in mind that the decision to move a student from one level of intensity to another should be a data-driven decision and, ideally, done in a consistent way across all implementation sites or at least within a school. The data that drives this decision should never be a single metric. If a single metric is used to make determinations about a student’s assignment to level of instruction, there is a higher likelihood for false positives (deciding that students need intervention, when they really do not) and false negatives (deciding that students do not need intervention, when they really do). When teams use multiple data sources for data-based decisions, they will make fewer errors in either direction (fewer false positives and false negatives), and more effectively use limited resources.
If the RTI model is working, there should be movement of students who were identified as in need of more intensive interventions (i.e., secondary or tertiary) to less intensive levels of instruction (e.g., student moves from tertiary to secondary or secondary to primary). This reflects movement in a desired direction.

Likewise, if the RTI model is working, there will be a subset of students who are identified as nonresponders (e.g., need more intensive interventions) and will move from primary to secondary or secondary to tertiary.

If the model is working, there should be an imbalance in this movement, or more students moving from tertiary to secondary to primary than moving from primary to secondary to tertiary.

It is important to remember that secondary prevention is not a long-term placement and that if students are stuck in the secondary level of prevention, it may suggest that there is a problem with the core instruction.
Let me illustrate the concept. If the RTI model is working, the number of students who move to less intensive levels of instruction would be greater than the number of students who move to more intensive levels of instruction.
If the RTI model is NOT working, the number of students who move to less intensive levels of instruction would be lower than the number of students who move to more intensive levels of instruction.
Ways to Look at Movement Between Levels of Prevention

1. Overall changes in the percentage of students within each level across the year
2. Percentage of students moving from a higher to lower level of intensity or lower to higher level of intensity
   a) Individual school
   b) Aggregated across schools

We are now going to look at two examples for how we might look at movement between levels of intensity. The first example looks at changes in the percentage of students within each prevention level from the fall to the spring. The second example compares changes in percentage of students moving from higher to lower levels of intensity and lower to higher levels of intensity for a single school or aggregated across schools.
For this analysis, you can also aggregate data across schools for an evaluation of RTI in general, not just at a specific site. This graph shows the percentage of kindergarten students receiving primary, secondary, and tertiary prevention levels in School A, a school implementing RTI. The data on student movement between levels indicate that this RTI model likely is effective.

Note: Students that are receiving secondary prevention should also be receiving primary prevention (core instruction). Decisions regarding student participation in both primary and tertiary levels of prevention are made on a case-by-case basis, according to student need. Tertiary level interventions need to address the general education curriculum in an appropriate manner for students (NCRTI Integrity Rubric).

The graph indicates that
- The percentage of students assigned to primary prevention in the fall (61%) is lower than the percentage of students in the spring (79%).
- 25% of the students were assigned to secondary prevention in the fall, whereas 15% of the students were assigned in the spring. This indicates a reduction in the percentage of students needing supplemental support.
- 14% of the students were assigned to tertiary in the fall, whereas 6% of the students were assigned to tertiary in the spring. This indicates a significant decrease in the number of students needing the most intensive support.

It is important to note that a decrease in the percentage of students in secondary prevention does not tell us whether they moved from secondary prevention to primary prevention (this would indicate that the RTI model is likely to be effective) or from secondary prevention to tertiary prevention (a large increase in this direction would suggest that the model is ineffective). In this example, we can see that the number of students in primary prevention increased and the number of students in tertiary prevention decreased, suggesting that students in secondary prevention moved back to primary prevention.

Note: All data in this presentation are fictitious.
It is also possible to look at movement between levels of intensity in more detail. For example, you may want to investigate the level of change between each level of intensity (e.g., primary to secondary, secondary to tertiary, tertiary to secondary). This graph shows the percentage of students who moved from a lower to a higher level of prevention and from a higher to a lower level of prevention across first-grade classrooms from fall to winter.

Remember, if the RTI model is working, there will be a subset of students who are identified as nonresponders (who need more intensive interventions) and will move from primary to secondary to tertiary. There also should be movement of students who were identified as in need of more intensive interventions (secondary, tertiary) to less intensive levels of instruction (e.g., tertiary to secondary, secondary to primary).

Results for movement from a less to a more intensive level of instruction show the following:
- 6.67% of the students within primary prevention needed the additional support of secondary prevention.
- 5.26% of the students who were in secondary prevention in the fall were not showing adequate progress and moved to tertiary prevention for more intensive, individualized instruction.

Results for movement from a more to a less intensive level of instruction, or our desired direction, show the following:
- 42.11% of the students who were in secondary prevention during the fall showed adequate progress and no longer needed to receive secondary prevention supports.
- 28.57% of the students in tertiary prevention in the fall moved from tertiary prevention to secondary prevention.
- 14.29% of the students in tertiary prevention in the fall moved from tertiary prevention to primary prevention.

The percentage of students moving from a more to less intensive level of instruction (84.97%) far exceeded the percentage of students moving from a less to more intensive level of instruction (11.93%), providing evidence of a positive effect for RTI.

Note: All data in this presentation are fictitious
This is a summary graph for the percentage of students who moved from a less to a more intensive instruction and from a more to a less intensive level of instruction aggregated across RTI implementation schools from fall to winter.

Group Question: What does this graph tell us?

Answer: Overall, the percentage of students moving from more to less intensive instruction (36%) exceeded the percentage of students moving from less to more intensive instruction (20%), indicating that the model's implementation is generally effective.

*Note: All data in this presentation are fictitious.*
Have participants turn to the Evaluation Plan Graphic Organizer and complete section 5, focused on movement between levels of prevention, with their teams. Provide 5–10 minutes for this activity.

Instructions: The handout includes sections focused on the considerations that we discussed in the previous section of the presentation. You will see questions related to data collection, staffing and expertise, funding, evaluation tools, timeframe, data analysis, and data sharing. With your team, answer the questions on the basis of what you are doing to evaluate movement of students between levels of prevention and what your plans are for the future.
Throughout the presentation, we have covered a lot of information. In order to review some of the key concepts and to come to consensus as a group on some information that will help to develop an evaluation plan, we are going to complete the **Guiding Questions for Evaluation Plan Development Handout**. The handout includes the following questions.

*Review the slide.*

### Guiding Questions Handout

- As a team, come to a consensus on the following guiding questions.
  - What do we want to know (e.g., evaluation questions)?
  - What is an indicator of what we want to know (e.g., success indicators)?
  - What type of data do we need to collect (e.g., format, existing or new)?
  - How will it be submitted or collected (e.g., by whom, data system use)?
  - How frequently will we collect this data? When will we collect and analyze?
  - Are there any potential challenges or concerns?
  - What fidelity data do we already have?
  - Do we need to adjust our process to make future evaluation efforts more successful? If so, what changes do we need to make?
Examples of answers

1a) You would look at progress monitoring data and examine whether students’ actual ROI is greater than the targeted and typical ROI.

1b) You could look at multiple forms of data (e.g., student outcome data or screening data) and compare the results for schools implementing RTI with those that are not.

1c) Any of the examples of data analysis that we discussed can be helpful in looking at whether your RTI model is working. It is important to look at multiple measures to have a more comprehensive understanding of whether your model is working. You may consider comparing the outcomes of your students on standardized tests or screening measures when you are implementing RTI to when you were not implementing RTI.

2. If we do not know whether we have implemented our RTI model and the essential components with fidelity it is impossible to know whether the outcomes we experience are a result of RTI or other variables. Measuring fidelity helps us to understand whether we have implemented RTI and the essential components as they were intended to be implemented.

3. Screening data, progress monitoring data, student outcome data.
Closing

- Good data in, good data out
- Evaluation plans help you understand whether your RTI model is working, and should include
  - Measures of fidelity
  - Measures of student, school, and district success
- There are multiple ways to evaluate your RTI model and analyze the data. Developing an evaluation plan is important.

Read slide.
Questions?

National Center on Response to Intervention

www.rti4success.org
National Center on Response to Intervention

This document was produced under U.S. Department of Education, Office of Special Education Programs Grant No. H326E070004. Grace Zamora Durán and Tina Diamond served as the OSEP project officers. The views expressed herein do not necessarily represent the positions or policies of the Department of Education. No official endorsement by the U.S. Department of Education of any product, commodity, service, or enterprise mentioned in this publication is intended or should be inferred. This product is in the public domain. Authorization to reproduce it in whole or in part is granted. While permission to reprint this publication is not necessary, the citation should be: www.rti4success.org.