"Selecting Evidence-Based Practices"

Slide 1: Welcome to the webinar, "Selecting Evidence-Based Practices." This is one of 11 webinars developed by the National Center on Response to Intervention (NCRTI). This webinar focuses on the steps needed to select evidence-based practices and provides guidance on how to utilize several online resources that provide information on curricula and interventions. My name is Jenny Scala, and I am one of the technical assistance providers at the Center. I have been working for the Center for a number of years in providing supports to a number of states on implementing Response to Intervention.

Slide 2: The Center has developed a series of webinars to cover information that is important for RTI implementation. On this slide you can see the different webinars available in the series. While you can watch the webinars in any sequence, we recommend that you first watch "What Is a Multi-level Prevention System?" since this webinar assumes a general understanding of the RTI multi-level prevention system.

Slide 3: During the presentation today, we'll first go through some steps for selecting evidence-based practices, followed by a demonstration for using online resources. We will be highlighting three websites: the What Works Clearinghouse, Best Evidence Encyclopedia, and NCRTI Instructional Intervention Tools Chart.

Slide 4: Upon completion of this webinar participants will be able to:

- Define characteristics of evidence-based interventions and research-based curricula;
- Use online resources to identify, evaluate, and select evidence-based interventions and practices; and
- Evaluate fidelity of implementation and outcomes.

Slide 5: One helpful way to check your understanding throughout the webinar is to complete the accompanying vocabulary handout. The handout provides a table with a list of key terms on the left side with columns to the right side for you to track your understanding before and after viewing the webinar. Before viewing the webinar, you should make predictions of the meanings of each term. Then, throughout the webinar, you will complete the final meaning column, based on the definitions provided, along with the picture/sketch/example column, where you can add anything helpful for remembering what you have learned. You can see how in this example, I was able to clarify the meaning of primary prevention level after watching the webinar.

If you have not already made predictions for the key terms of this webinar, please go ahead and pause the webinar so that you can do so now. Press play when you are ready to continue.

Slide 6: NCRTI refers to an evidence-based intervention as an intervention for which data from scientific, rigorous research designs have demonstrated (or empirically validated) the efficacy of the intervention. That is, within the context of a group or single-subject experiment or a high-quality quasi-experimental study, the intervention is shown to improve the results for students who receive the intervention. Research-based curricula, on the other hand, may incorporate design features that have been researched generally without the curriculum or the program as a whole having been studied using a rigorous research design as identified and defined by the Elementary and Secondary Education Act.

Slide 7: As a reminder, NCRTI recommends different evidence standards across intervention levels. The research-based curriculum is recommended for primary prevention across subjects. This means, again, that components have been researched and found to be generally effective. However, the curriculum materials as a whole have not been rigorously evaluated as a package

Slide 8: Evidence-based interventions are recommended for students who are receiving support at the secondary and tertiary levels of prevention. This means that the curriculum has been evaluated using a rigorous research design and that there is evidence of positive effects for students who received the intervention.

Slide 9: Selecting evidence-based practices involves a multistep process that demands the consideration of needs and priorities and the identification of practices that match those needs and priorities. A critical and sometimes overlooked step in this process is a review of existing evidence to ensure that identified practices meet established criteria for being evidence based. Once evidence-based practices are selected, the process shifts to implementing those practices with fidelity and evaluating the effectiveness of those practices in improving student outcomes. To select appropriate evidence-based practices, the following steps are recommended.

1. Identify needs and priorities.
2. Select practices to address needs.
3. Evaluate evidence claims.
4. Implement practices.
5. Evaluate effectiveness.

The first step recommended in selecting an evidence-based practice is to identify needs and priorities.

Slide 10: In identifying needs and priorities you should:

1. Gather a team. The team should include key stakeholders in the school or district. The team's composition may vary, but it is recommended that the team include a diverse group of stakeholders. At least one member should be familiar with data interpretation and evaluating research claims.

2. Conduct a needs assessment
   - This allows you to gather information from multiple sources—for example:
     - Teachers
     - Parents
     - Administrators
     - Students
     - Community members
   - You then compile data to identify areas of need—for example:
     - Academics
     - Behavior
     - School climate
     - Professional development

3. Determine priorities or what factors they consider most important.

There is more information found in the National Center's Instructional Intervention Tool Chart User's Guide.

Slide 11: In determining needs and priorities it is important to look at the baseline performance of indicators of interest. This might include the academic achievement of students, disciplinary referrals, or attendance and truancy data.

Slide 12: In thinking about your needs, ask yourself and your team these types of questions:

- For what skills do we need a secondary intervention instructional program? Is there a specific academic outcome or measure for which we are interested in providing supplemental instruction?
- For what grades do we need an instructional program?
- Will this program be used with all students who are not progressing in the core curriculum or only with specific subgroups of students?
- Which subgroups?

Again, there is more information in the *Instructional Intervention Tool Chart User's Guide* produced by the Center.

Slide 13: In addition to determining your needs for an instructional program or an intervention, your team should also consider its priorities. Although you may want a program that meets all of these criteria, such a program may not exist. You will need to weigh your priorities carefully when making your decision. Often teams overlook these steps, so it is really important that you discuss these questions in order to determine your priorities.

Slide 14: After teams have identified their needs and priorities, they will need to identify practices that address those needs and priorities.

Slide 15: Evidence-based practices are not only essential for multi-level instruction. As discussed in previous modules and webinars, it is important to identify evidence-based screening and progress monitoring tools. Data collected from these tools should be interpreted using evidence-based decision-making criteria. We can then use this data to select evidence-based interventions that then match student needs. The next few slides provide a couple of examples of resources where you can find evidence-based practices. See Appendix C in the training manual for descriptions and links to more resources.

Slide 16: As we learned from previous trainings, evidence-based progress monitoring and screening tools can be found using a variety of resources. Here are just a few:

- NCRTI: Screening Tools Chart
- NCRTI: Progress Monitoring Tools Chart
- IES Practice Guide: RTI for mathematics and reading
- IRIS Center modules and RTI/assessment training modules

Slide 17: We have also discussed several resources available to assist teams in identifying evidence-based data decision criteria.

- IES Practice Guide: Using Student Achievement Data
- U.S. Department of Education: Doing What Works and the American Recovery and Reinvestment Act
- IRIS Center Case Study Unit RTI: Data-Based Decision Making

Slide 18: Numerous resources exist for identifying evidence-based practices. The first three resources, What Works Clearinghouse (WWC), Best Evidence Encyclopedia (BEE), and the NCRTI Instructional Intervention Tools Chart, are searchable databases that allow users to search for interventions by a number of features (e.g., grade level, content, and delivery method). Although these sites are similar, each has its own criteria for determining what is evidence based. These will be discussed in more detail later in the presentation. It is important to be aware that understanding these criteria is critical as you are using the different websites. The Practice Guides are published documents from the Institute of Education Sciences (IES), which provide summaries of research-based practices in math and reading. The IRIS Center (IDEA and Research for Inclusive Settings Center) modules provide training and research on various teaching practices.

Slide 19: Once a team has identified practices that meet its needs and priorities, the next step is to evaluate the evidence claims.

Slide 20: When teams evaluate evidence, they should consider the following questions:

- Where can I find evidence?
- What type of evidence exists?
- What is the quality of the evidence?
- What were the desired outcomes?
- What are the effects of the intervention?
- Is the sample population similar to my own population?

We will discuss these questions in more detail in the next slides.

Slide 21: Evidence for programs can be found on curriculum websites and in various peer-reviewed journals (which can be found through a number of sources including ERIC [Education Resources Information Center] (www.eric.ed.gov/), Google Scholar (www.scholar.google.com/), Education Abstracts, Psychological Abstracts, and more). If these resources are used, it is important for the team to have established review criteria prior to evaluating the evidence. Caution should be taken when evaluating evidence provided by a publisher. Some questions to consider are the following:
- Did an independent researcher conduct the study?
- Are the results from a rigorous research study?
- Is the study population similar to the school or district population of interest?
- What outcomes and skills were targeted?

To assist educators, several resources publish independent reviews of instructional practices using rigorous review criteria. The outcomes of these reviews can be found in the IES Practice Guides, What Works Clearinghouse, Best Evidence Encyclopedia, and the NCRTI Instructional Interventions Tools Chart. Links to these resources can be found in Appendix C of the Training Manual.

Slide 22: Because access to peer-reviewed journals, access to established review criteria, and the time to sift through existing research are limited for most teams, several groups (funded through federal funds or state funds) have done this work. The following are examples of sites that have conducted rigorous reviews of existing educational practices or interventions and then published the results. It is important to remember that each site has its own review criteria and expected level of rigor. For example, the What Works Clearinghouse has the most stringent review criteria, which results in fewer practices being considered evidence based. You might find that a tool is considered evidence based on one site but not another. You may also find that some tools are on some sites but not on others. This is because there are also differences in how tools or practices are selected for review. For example, if a publisher says it has an intervention, then the What Works Clearinghouse conducts a thorough review of peer-reviewed journals to identify all existing evidence and then rates the program based on the reviewed evidence. The NCRTI

Instructional Intervention Tools Chart, on the other hand, reviews the research of tools submitted by the publisher. As a result, NCRTI does not rate the program; it rates the submitted research evidence. Regardless of the site used to identify evidence-based programs, it is important to understand how the tools were selected and what criteria were used to rate them.

Slide 23: Evidence may come in different formats and from different sources. Teams can review actual research studies, such as a technical report that provides evidence about its effectiveness, or a summary of the existing research, which is common on the websites we just discussed as well as in the IES Practice Guides. Many people often depend on peer or publisher claims, such as a district that believes a program is effective because a neighboring district claimed it raised test scores. Teams may find other evidence, such as a brief claiming the positive effects of research conducted by the publisher or research conducted within the school district. It is important to review these types of reports closely to determine the quality of the evidence, which we will discuss in more detail.

Slide 24: Evidence not only takes different forms, but also includes various types of research studies (e.g., randomized control trial, single-case designs, quasi-experiments, and quantitative research synthesis).

- While not always feasible, randomized control trials are considered the most rigorous form of research. Eligible subjects are randomly assigned into groups to receive or not receive one or more interventions that are being compared.
- Carefully designed studies with quasi-experimental designs, which do not randomly assign students, attempt to best approximate the effects of randomization that are found in randomized control trials through a number of different methods to minimize selectivity bias. Both types of studies work to ensure that the effects shown for the intervention are not based on extraneous variables or initial differences between the control and experimental group but are a result of the intervention.
- Single-case designs are research studies where the researcher attempts to demonstrate an experimental treatment effect using single participants.
- Quantitative research synthesis uses statistical techniques to combine multiple quantitative research studies. An example is a meta-analysis.

Slide 25: In addition to looking at where the evidence came from and the type of evidence provided, it is important to look at the quality of the information provided and whether it is feasible to implement the program under normal conditions. Some questions to consider in looking at the information provided are as follows:

- Who implemented the intervention?
- Was the intervention described?
- Was there a manual or a script?
- How often did the intervention occur?
- Was fidelity evaluated in the study? If so, how?

- What phase of research? (exploratory, pilot, efficacy, scale-up)

It is also important to consider how feasible it would be to replicate what was done in a classroom in a school or district. The following should be considered when thinking about the feasibility of the intervention:

- How complex is the intervention?
- How much staff training will be necessary to implement it?
- How often should the intervention be delivered?
- What was the student-teacher ratio?
- How much does it cost (materials, training, staff time)?

Slide 26: It is also important to think about what outcomes were measured in the study and whether the outcomes assessed are relevant to the outcomes of interest and are feasible for your population. In addition, it is important to know what measures were used to evaluate the intervention outcome and if these measures are reasonable, relevant, reliable, and valid. The measures might include both proximal measures, which assess aspects of competence that the program was directly targeted to improve, and distal measures, which assess aspects of competence that are related to the skills targeted by the program but not directly taught in the program.

Slide 27: After considering the rigor of the study, the next step is to evaluate the effects. Effect sizes are useful in comparing programs and identifying programs with greater potential for student gains. The effect size is a standardized measure of the magnitude of the relationship between two variables, or more specifically, in this case, the magnitude of the relationship between the performance of students participating in a particular intervention and an outcome of interest. The larger the effect size, the larger the impact participating in the intervention had on the outcome. Generally, research literature identifies effect sizes of 0.80 or greater as "large" effects, 0.50 as "moderate" effects, and 0.20 to 0.30 as "small" effects (Cohen, 1988).

You want to look at programs that not only are supported by evidence from technically rigorous studies, but also show positive program results. Various sites report the effectiveness of programs differently. For example, the NCRTI Instructional Intervention Tools Chart calculates the effect size across all studies submitted, using a standard formula, and then offers this information so that viewers of the chart can compare for themselves across studies for the outcome measures in which they are interested. The What Works Clearinghouse presents this information through what they refer to as an improvement index. Some sites fail to show the strength of the effect and instead just rate the program as effective or not. It is recommended that effect size be reviewed in making decision about programs. It allows teams to make comparisons and determine which programs are likely to result in the level of growth needed to close the achievement gap.

Slide 28: It is important to review the sample population in the study to determine whether the intervention has been evaluated for students similar to those in the district, school, or population of interest. For example, teams may be interested in interventions with positive effects for targeted groups of students, such as students with low socioeconomic status, English language learners (ELLs), or students in special education. Some sites that summarize program research do not report this level of information. Teams may consider using multiple resources to evaluate the existing evidence of a particular program.

Slide 29: Now that you understand what to look for, let's take some time to see how several resources can assist educators in making decisions about selecting evidence-based practices. We will be focusing on these:

- What Works Clearinghouse
- Best Evidence Encyclopedia
- NCRTI Instructional Intervention Tools Chart

All of these resources are online and are accessible to anyone. At this point in the presentation, I will provide an overview of the resources available on each website. You have two options. You can continue to listen to the presentation while you explore the websites as I talk about them. Or, you can continue watching the slides, which are screen shots of the websites, and then go back and visit them at a later time. Remember that if you need more time to look at a website, you can always pause the webinar before moving on.

Slide 30: The first resource we'll be discussing is the What Works Clearinghouse, which includes rigorous reviews of programs available in a variety of content areas. If you wish to look at the website now, open a new tab in your browser and type in "http://ies.ed.gov/ncee/wwc/". To find programs, you can click on the top tab labeled "Topics" or click on a specific topic under the tab, or you can click on the Topics box located on the right side or one of the specific topics listed there. For now, let's click on the Topics tab at the top (*general page*). Knowing your needs and priorities prior to using the WWC is essential for selecting appropriate interventions in an efficient manner.

Slide 31: On the general topic page, you will notice that it says, "Topics are defined by intended outcome (such as improved math skills), intended population (such as students with special needs), and types of interventions (such as school choice) that may improve student outcomes." If you would like to look at topics by intended outcome, you should click on the drop-down menu below "Find WWC Topics." Also notice that the topics by intended population and type of intervention are listed below. For each of these topics, you'll see a one-sentence description of the topic, followed by a bulleted list of outcomes reviewed, and a list of latest publications. You'll also notice that the green box on the left side says, *"Not sure where to start? We can walk you through it."* This tab provides additional tips about how you can use the WWC webpage. For now, let's click on the green box on the right side of the screen that says,

"Find what works and create your own report." This will provide us with a more extensive search feature.

Slide 32: Once in the search feature, you can search by the name of the intervention or select your needs and priorities. Again, you'll see the outcome domains on the left. For desired outcome domains, let's choose reading achievement. For grade levels, click on second grade. For population, click on "general education."

Slide 33: You can also select the level of effectiveness, extent of evidence, delivery method, and program type. For our search, let's not click anything for effectiveness, extent of evidence, or delivery method. Let's pick curriculum for the program type.

It is important to understand that only interventions with research evidence that meet the WWC standards are included in the search. The number of interventions that meet the criteria for that search feature is included in parentheses.

Now that we have selected our search features, we can create a report by clicking Find.

Slide 34: Now we can see the programs that meet our criteria. You can see the names of the interventions, the topic areas of focus, the improvement index, the effectiveness rating, and the extent of the evidence. Above the results of your search is a drop-down menu that says "Sort by," which allows you to sort by intervention name, improvement index, rating of effectiveness, and extent of the evidence. If you click on the tabs, you can learn more about the intervention details and research details. Let's click on Intervention Details.

Slide 35: Let's look more specifically at the Accelerated Reader program. Click on Intervention Details to learn more about this program.

Slide 36: The next step is to evaluate the program to see if it meets our needs and priorities. What Works Clearinghouse does not rank or recommend programs. Instead, it provides data to support teams in the selection process. This is why it is critical that teams clearly identify the needs of their at-risk population. For example, the first program listed, Accelerated Reader, appears to be more effective for reading achievement than reading fluency or comprehension. This is a mismatch between student needs and the outcomes of the intervention. A mismatch between student needs and the outcomes of the intervention could result in nonresponsiveness and be misinterpreted as a student skill deficit.

Slide 37: The improvement index determines the difference between the percentile rank of the average student in the intervention group and the percentile rank of the average student in the comparison group—in other words, the effects. The improvement index can take on values between -50 and +50, with positive numbers denoting results favorable to the intervention group.

The smaller the index, the smaller the gains students made from the comparison group. For example, in this case the differences between students in the Accelerated Reader intervention and those in the comparison group are minimal for reading comprehension outcomes. There was a greater difference in reading achievement than reading fluency or reading comprehension. A negative index would indicate students made less progress in the program than other students. Scores closer to 50 are considered to demonstrate greater gains.

Slide 38: For each intervention, the effectiveness rating summarizes the results of the studies that met WWC evidence standards. The effectiveness rating shows whether the studies indicate that the intervention has positive effects, potentially positive effects, mixed effects, potentially negative effects, negative effects, or no discernible effects. The rating also reflects the degree of confidence in the estimated effects. The effectiveness rating takes into account four factors: the quality of the research design, the statistical significance of the findings, the size of the difference between participants in the intervention and the comparison conditions, and the consistency in findings across studies.

If you click on the green button that says "Glossary" on the top right of your screen, you can scroll down to view the Effectiveness Rating Key to see what each symbol means. The effectiveness rating has six categories:

1. Positive Effects*:* Indicates strong evidence of positive effects with no overriding contrary evidence.
2. Potentially Positive Effects*:* Indicates evidence of positive effects with no overriding contrary evidence.
3. Mixed Effects: Indicates evidence of inconsistent effects. At least one study shows a statistically significant or substantively important positive effect. At least one study shows a statistically significant or substantively important negative effect, but there are fewer studies that show statistically significant or substantively important negative effect than show a statistically significant or substantively important positive effect.
4. No Discernible Effects*:* Indicates no affirmative evidence of effects. None of the studies shows a statistically significant or substantively important effect, either positive or negative.
5. Potentially Negative Effects*:* Indicates evidence of negative effects with no overriding contrary evidence.
6. Negative Effects*:* Indicates strong evidence of negative effects with no overriding contrary evidence. Two or more studies show statistically significant negative effects, at least one of which is based on a strong design. No studies show statistically significant or substantively important positive effects.

Slide 39: The extent of evidence category was developed to inform readers about how much evidence was used to determine the intervention evidence rating. It focuses on the number and the size of studies. This scheme has two categories: small and medium to large. Domains with a small extent of evidence include only one study, only one school, or findings based on a total

sample size of less than 350 students; assuming 25 students in a class, there are fewer than 14 classrooms across studies. Domains with a medium to large extent of evidence include more than one study, more than one school, and findings based on a total sample size of 350 students; assuming 25 students in a class, a total of at least 14 classrooms across studies.

Slide 40: To evaluate the evidence in more detail and its relevance to your needs, click on either the name of the program at the top in green or on the words "Read full report."

Slide 41: The report includes information about the number of studies identified and that met the criteria. For example, for this program, 100 studies were identified, but only two met WWC evidence standards. The remaining studies did not meet WWC evidence screens. This report also includes cost, population samples, publisher building, description of the intervention, and the range of effects. This page shows that although the average for general reading comprehension was 0 percentile points, the range was -12 (meaning some students performed 12 percentile points lower) to +12 (scored 12 percentile points higher).

Slide 42: Another useful site is the Best Evidence Encyclopedia. If you are able, open a new tab in your browser and go to http://www.bestevidence.org/ to follow along. The Best Evidence Encyclopedia uses less stringent review criteria than WWC. At the top of the page, you'll notice that there is a menu that includes "About the BEE" and "Review Methods." Before using this site, it is recommended that you review this information to understand the criteria for inclusion of programs and how they are reviewed.

On the left side of the homepage, you will notice the Program Reviews list. Here you can click on the relevant topic to review programs and their reviews. Knowing your needs and priorities prior to using the Best Evidence Encyclopedia is essential for selecting appropriate interventions in an efficient manner. For this demonstration, let's select middle and high school mathematics.

Note: These slides represent the information provided on the website as of December 2011.

Slide 43: At the top of the page, you can find programs that show high or moderate effects, find programs that show limited effects, find programs with insufficient evidence, and see key findings, which is a summary of the review and the methods used to review the programs for middle and high school mathematics.

Slide 44: Below these options, we find a short summary of the research report. The information in this section is based on a review that summarizes evidence on three types of programs designed to improve the mathematics achievement of students in Grades 6-12:

- Mathematics Curricula (MC), such as *The University of Chicago School Mathematics Project, Connected Mathematics, Saxon Math,* and other standard and alternative textbooks.

- Computer-Assisted Instruction (CAI), such as *I Can Learn, Jostens/Compass Learning,* and *Accelerated Math.*
- Instructional Process Programs (IP), such as cooperative learning, mastery learning, and other approaches primarily intended to change teachers' instructional strategies rather than curriculum or technology.

The full report or an educator's summary can be accessed by clicking on the link on the right side.

Slide 45: Let's now go back to the top menu and select Review Methods. After you click on this, you can see the review criteria used to determine a program's effectiveness. Based on the effect size, the type of study, and the size of studies reviewed, programs receive one of five ratings: strong evidence, moderate evidence, limited evidence, insufficient evidence, or no qualifying studies. Scroll down to see the definitions of the ratings.

Slide 46: Now, let's go back up to the top menu and select Top Rated Programs. You'll notice that the report shows that two middle and high school programs were identified as having strong evidence, and no programs were identified as having moderate evidence. In other words, of all the programs they reviewed only two met their criteria as an evidence-based program for middle and high school mathematics. Let's take a closer look at the information the report provides in this table.

Slide 47: In the first column, an indicator (rating) of effectiveness is provided. In this case, this program has strong evidence of effectiveness, which means there were at least two studies that met the criteria that showed an effect size of at least .20. Remember, earlier I mentioned that the research literature identifies effect sizes of 0.80 or greater as "large" effects, 0.50 as "moderate" effects, and 0.20 to 0.30 as "small" effects. The actual effect size is not provided in this chart.

Slide 48: The third column indicates the type of program. Remember, the WWC reported this differently—curriculum, supplemental, or practice. For this example, the program is an Instruction Process Program (IP), such as cooperative learning, mastery learning, and other approaches primarily intended to change teachers' instructional strategies rather than curriculum or technology.

Slide 49: The remaining information is about the program, including its name, a brief description, and contact information.

Slide 50: In 2010, the Center released its Instructional Intervention Tools Chart to help users determine the effectiveness of interventions used in secondary prevention systems. You can go to the tools chart now by typing www.rti4success.org/instructiontools. The first column has the program name and the second column shows the research study, or studies, evaluated for that particular program. Indicators of study quality and effect size data are provided in the remaining columns to assist consumers in evaluating the efficacy of the specific program. All programs on

the Tools Chart have been submitted for review. In other words, it is voluntary for programs to send their information to the National Center's review committee.

Slide 51: As I just mentioned, each tool on these charts is presented because the developers voluntarily submitted them to NCRTI for review in response to one of NCRTI's annual "calls for submissions." That means that the Instructional Intervention Tools Chart is by no means an exhaustive list of every available intervention. It's important to keep this in mind: these charts are intended to be a source of information that can help users select tools and programs, and they can help you think about what information you need to know to select an appropriate tool or a program. If there is a tool or a program that you are considering using but it is not on the chart, we recommend that you call the developer directly and ask them to submit their product to our review process.

Unlike the previous resources, the NCRTI Instructional Intervention Tools Chart does not provide an overall rating for the program. Instead, it reviews the supporting research that is specific to RTI.

Slide 52: For a tool to be reviewed, it first had to first meet the instruction definition. This is the operational definition of instruction that was developed by NCRTI's Technical Review Committee (TRC).

NCRTI definition of instruction: Additional or alternative intervention programs to the core curriculum conducted in small groups or individually with evidence of efficacy for improving academic outcomes for students whose performance is unsatisfactory in the core program.

This definition of instruction reflects instruction that would occur in what we would typically refer to as secondary intervention.

Slide 53: The goal of the Instruction TRC is to identify those standard treatment protocols that have been found through research to be effective: that is, to help users identify programs that have, first, been studied through rigorous design, and second, have shown positive, meaningful treatment effects. The Center in no way recommends any specific programs, nor are the programs on the chart endorsed more than those that are not on the chart.

Slide 54: At the top of the Tools Chart, you'll notice that there are two drop-down options. This is a new feature, which allows you to sort tools and programs by the subject (math and reading) and grade (elementary and secondary). To use these features select the appropriate subject or grade level and press the Filter button. If you want to return to all the tools simply press Reset.

Slide 55: The tools chart also allows you to narrow your search by selecting certain interventions or studies to compare side by side. To select certain interventions or studies, simply check the box in the last column of the chart. Go ahead and select as many as you would like to compare.

Because the Instructional Intervention Tools Chart looks at the research studies of interventions, you may have more than one study for the same intervention. For example, when looking at Corrective Reading Decoding, you'll see that there are multiple research studies. When looking at the Tools Chart you may want to compare different interventions or you may want to look at the different studies available for a single intervention. The compare feature allows you to do either of these options.

Slide 56: Now press the compare button and the chart will appear with the interventions or studies you selected for side-by-side comparison.

Slide 57: The Instructional Intervention Tools Chart includes information on three aspects of a program: study quality, effect size, and implementation requirements. You can click on the program name to see the implementation requirements. Click on a program of your choice in the first column of the chart, and this information will appear.

Slide 58: A pop-up will come up with information about the intervention, usage requirements, the training needed for the intervention, the cost of the intervention, and the program specifications and requirements.

Slide 59: You can click on the links to find more information about the participants, the study design, fidelity of implementation, and measures.

The TRC used the following questions to judge these components when reviewing them:

- Participants. Are the students in the study at risk? Are the program instructors in the study similar to what the vendors state is necessary?
- Design. Does the study design allow us to conclude that the intervention program, rather than extraneous variables, were responsible for the results?
- Fidelity of Implementation. Was it clear that the intervention program was implemented as it was designed to be used?
- Measures. Were the study measures accurate and important? The chart provides information for both proximal and distal measures. Remember, earlier we defined proximal measures as measures that assess aspects of competence that the program was directly targeted to improve, and distal measures are those that assess aspects of competence that are related to the skills targeted by the program but not directly taught in the program.

Slide 60: For participants, the TRC looks for evidence that the sample that was studied is indeed considered "at risk." Because these are secondary interventions, they are designed to target the needs of at-risk students, so the sample should have those qualifications. Users can click on the bubbles to find the information used to rate the tool. This information also allows the users to compare their population to the populations researched in the study. Click on a circle in the

participants column and a pop-up will provide information about the sample size, the inclusion of ELL students, and grade levels of the sample.

Slide 61: For design, the TRC generally is looking for studies that have been designed in such a way that any results can be attributed solely to the intervention, not to any extraneous factors. This means that the study (ideally) used random assignment or it was a high-quality quasi-experimental design, which means that any systematic differences between the control group and the treatment groups are small and are accounted for appropriately in the analyses.

The following elements were used to determine the rating:

- Did the study use random assignment?
- If not, was it a tenable quasi-experiment?
- If the study used random assignment… You'll see that this slide shows in the pop-up the additional questions that were used.

Slide 62: The next column shows rates of fidelity of implementation, which means that the study was able to document how the intervention was implemented and whether or not it was implemented in exactly the same way as it was intended.

Slide 63: In the next two columns, measures, the TRC looks at whether the outcome measures that were assessed in the study are relevant; in other words, first, are the outcomes related to a program's instructional content and, second, have these measures been proven to be reliable and valid? The TRC also wants to see studies that look at a range of both proximal and distal measures.

When you click on either the proximal or distal measure bubble, you will obtain the measure, the score type, the range of measures, the reliability statistics, and the relevance to the program instructional context.

Proximal measures assess aspects of competence the program was directly targeted to improve. Typically, this does not mean the very items taught but rather novel items structured similarly to the content addressed in the program. For example, if a program teaches word attack, a proximal measure would be the decoding of pseudowords. If a program teaches the comprehension of cause-effect passages, a proximal measure would be answering questions about cause-and-effect passages structured similarly to those used during intervention but not including the very passages used for intervention. In this example Woodcock Johnson–III Basic Reading Skills, Woodcock Johnson–III Letter-Word Identification, Woodcock Johnson–III Word Attack, and DIBELS Oral Reading Fluency were all proximal measures.

Distal measures assess aspects of competence that are related to the skills targeted by the program but not directly taught in the program. For example, if a program teaches word-level reading skills, a distal measure would be answering questions about passages the student reads. If

a program teaches calculation skills, a distal measure would be solving word problems that require the same kinds of calculation skills taught in the program. In this example there were no applicable distal measures.

Slide 64: To the right of the study quality columns are the effect size columns. There are many different methods for calculating effect size. In order to ensure comparability of effect size across studies on this chart, NCRTI used two standard formulas to calculate effect size across all studies and outcome measures. Developers of programs on the chart were asked to submit the necessary data to compute the effect sizes. Results from both methods are reported on the chart. Effect sizes are separately reported for both proximal measures and distal measures.

The adjusted posttest refers to posttests that have been adjusted to correct for any pretest differences between the program and control groups.

Slide 65: The unadjusted posttest does not account for differences in the program and control groups. It is important to note that the unadjusted posttest is considered less rigorous than the adjusted posttest, because means have not been adjusted to account for pretest differences. This formula is typically used only in instances where we can assume pretest group equivalency.

Slide 66: You'll notice that many of the studies have a "—" in the effect size cell. These studies either do not have the necessary data or do not meet the assumptions required for calculating and reporting effect size using the associated formula. The reason for the missing data is provided when you click on the cell.

Slide 67: Let's look at an example. Click on the reset button on the right side if you have not already done so. Now scroll down to Number Rockets. In this example, the adjusted posttest for the proximal measures was 0.45 and for the distal measures was 0.10. Remember, earlier we mentioned that the research literature identifies effect sizes of 0.80 or greater as "large" effects, 0.50 as "moderate" effects, and 0.20 to 0.30 as "small" effects.

Slide 68: Now click on the cell in the "mean based on adjusted posttests" column. The pop-up shows that the effect of the Number Rocket intervention varied based on the outcome measures (or outcomes). This intervention had little effect on the participants' applied problems on the Woodcock Johnson Applied Problem and fact retrieval for subtraction *(distal measures)*. It had moderate effects on story problems, first-grade concepts and applications, and Woodcock Johnson Calculations. This information can be useful in finding interventions that match the desired outcomes.

Effect sizes should only be compared when they are calculated using the same method (e.g., effect size based on adjusted means cannot be compared with effect size based on unadjusted means).

Remember, it is important to review study quality and effect size together. For example, you need to make sure that a study that shows large effects does not show low ratings for quality, because you are then not confident that the effects shown are a result of the intervention.

Slide 69: In the last column, the tools chart reports effect size data that have been disaggregated for subgroups, if available—for example,

- Students with disabilities
- ELLs
- Students from diverse racial-ethnic groups.

In the Number Rockets example, there are no available data.

Slide 70: Just selecting an evidence based intervention is insufficient for increases in student performance. The intervention must then be implemented as designed.

Slide 71: Implementation involves providing initial recommended training and professional development to the interventionist. A plan must be developed and put in place to ensure that the intervention is delivered as designed, including the appropriate time scheduled and the appropriate materials. After the initial implementation begins, it is important to provide ongoing coaching and professional development to ensure that the intervention continues to be delivered as originally designed. Ongoing monitoring of implementation fidelity is essential.

Slide 72: We hear a lot about fidelity of implementation when asked about RTI. What does this really mean?

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.

Slide 73: Self-report data can include questionnaires, surveys, or interviews, and may provide an indicator of teacher knowledge as well as the context of implementation. However, it is important to realize that these measures are often unreliable when used alone.

Slide 74: 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.

Slide 75: Reviewing logs/lesson plans and student work allows for evaluation of what was done. It could include looking at the content covered and student progress. It provides less information about delivery, dosage, and adherence to scripts (if applicable).

Slide 76: Using Fidelity Data

It is important to use these data to distinguish curriculum/intervention from the quality of implementation when problems occur.

Slide 77: When we are evaluating our implementation and using fidelity data, here are some things to keep in mind:

- We can identify implementation strengths in terms of people, potential coaches, and components of intervention.
- We also want to make sure that we are targeting areas in need of improvement. This also applies to coaching, professional development, and possible retraining as needed.

Slide 78: The final step in selecting evidence-based practices is to evaluate our effectiveness. Once the intervention is implemented with fidelity, teams can evaluate the effectiveness of the intervention for groups and individual students.

Slide 79: Evaluating the effectiveness of the instruction cannot occur until the intervention has been delivered with fidelity. Otherwise, it is difficult to determine whether nonresponsiveness is a result of poor implementation or student learning differences. Before determining individual student responsiveness to an intervention, the effectiveness of the intervention for most students must be determined. If most students are not benefiting (and data indicate high fidelity of implementation), the intervention may not be effective for your student population. Once the effectiveness of the intervention for most students is determined (data indicate that the intervention is effective) it is easier to identify nonresponders that are a result of learning differences, not a result of a poor intervention.

Evaluating the effectiveness for individual students can be done by using progress monitoring data that show trend lines. This is one of the easier ways to show growth over time. You can also monitor groups of students by averaging weekly data points to show a group trend. In thinking about this, remember that

- Ongoing, graphed progress monitoring data illustrate progress over time.
- Data should be reviewed at least every 4–6 weeks. Remember from the previous training that it takes at least eight data points to create a trend line.
- Decisions are based on evidence-based decision criteria. Criteria for making decisions about tier movement and responsiveness should be based on pre-established, written criteria using logical practices.
- General outcome measure vs. mastery measure. As mentioned in previous trainings, it is important to distinguish whether the progress monitoring measure is a general outcome measure (one that measures an overall curriculum and is curriculum independent) or a mastery measure (a measure that is assessing sub-skills). By using the mastery measures,

we can determine if a student was able to acquire a specific skill. However, a general outcome measure will help us determine if a student is able to apply and maintain the skill.

Slide 80: As mentioned earlier in the presentation, practitioners can use progress monitoring data to compare and contrast the efficacy of interventions. The continued use of ineffective interventions can be costly and harmful to students. In this graph, intervention A appears to be more effective than B or C. Instead of providing three different interventions to address similar issues, it might be better to focus on intervention A. These types of data can also provide information about what is a realistic growth rate for students in these interventions. To obtain these data, average students' weekly scores to see growth of the group over time.

Slide 81: Once the intervention has been determined to be effective for most students, it is possible to begin identifying individual students who are not responding to the program. In this case, the data indicated that the student was not responding to an intervention that group data indicated was effective for most of the student's peers. The teacher decided to make a change to accommodate the student's learning difference (duration, frequency, intensity, or setting, etc.) and the data indicate the student is responding positively to the change.

Slide 82: Things to remember:

- Good data IN is going to give you Good data OUT.
  - Know where your data came from and the validity of that data. This also includes fidelity of implementation data.
- Focus on the big picture or ALL students.
  - Are most students making progress? We want to move away from looking at just individual students. We want to make sure that we can keep the individual student's growth in perspective of the bigger picture of how all students are doing.
- ALL instructional and curriculum decisions should be based on DATA.
- Finally, keep it SIMPLE and EFFICIENT.

Slide 83: If you need more information, please make sure that you look at these following websites:

- National Center on Response to Intervention: www.rti4success.org
- RTI Action Network: www.rtinetwork.org
- IDEA Partnership: www.ideapartnership.org

Thanks for taking the time to listen to "Selecting Evidence-Based Practices." To find more resources on this topic or to view other webinars in the implementer series, visit www.rti4success.org. You can also find more information from the RTI Action Network and IDEA Partnership sites.