

## **Applying the What Works Clearinghouse Standards to Postsecondary Research Webinar Transcript**

*Chris:*

Thank you for joining us. I am Chris Weiss, Senior Education Research Scientist at the federal Department of Education's National Center for Education Evaluation and Regional Assistance.

I'd like to welcome you to this What Works Clearinghouse webinar on applying WWC Standards to Postsecondary Research.

We are pleased to have a growing set of archived webinars and other resources that we hope will be helpful to you in designing, carrying out, analyzing, and interpreting your results – and the results of other studies in education. Later on in the webinar we will provide links and screenshots to these resources. We invite you to provide feedback on this webinar, on the available WWC resources, or on other ways we can better serve you as consumers.

In fact, the idea for this webinar came from some of you, and aims to respond to many questions we have received through the WWC help desk about our standards and procedures for reviewing studies, specifically those that focus on the postsecondary context. Our hope is that our conversations about the WWC's standards and procedures will aid the postsecondary research community and, especially, provide a helpful resource for future research in postsecondary contexts.

And with that, I'll turn it over to Jeff, Will, and Sandra.

*Jeff speaking:*

Thank you, Chris, and thank all of you for joining us today.

I'm Jeff Valentine, Professor and Program Coordinator of the Educational Psychology, Measurement, and Evaluation program at the University of Louisville and I am the Principal Investigator of the WWC's Postsecondary work with Development Services Group in Bethesda.

Will, can you introduce yourself?

*Will speaking:*

Sure! Hi, I'm Will Doyle, and I am an Associate Professor at Vanderbilt University's Peabody College and the content expert for the WWC's postsecondary topic areas. I'm joined by Sandra Wilson– Sandra, can you say hello?

*Sandra speaking:*

Hello, I'm Sandra Jo Wilson, Research Assistant Professor in the Department of Special Education and the Associate Director of the Peabody Research Institute at Vanderbilt University. Thank you all for joining us today.

*Jeff speaking:*

Great! We are really excited you all could join us to discuss the WWC's standards in the context of research on postsecondary education.

However before we get started we have some housekeeping to attend to.

First, the slide deck, a recording of the webinar, and a transcript of the webinar will be available on the WWC's website for download.

I also wanted to mention that you'll see that we have several hyperlinks embedded in the presentation. These hyperlinks won't work for the GotoWebinar presentation, but they will be active in the downloaded materials.

The next thing to mention is that we take questions! You can submit questions throughout the webinar, using the "chat" tool on the webinar software on your screen. We'll try to answer as many questions as we can at the end of the presentation. If we don't make it through all of the questions we've received, we will respond to your question via e-mail. And, all questions and responses will be available for download some time after the webinar ends.

Finally, we've scheduled about an hour for this webinar.

So, with that introduction, let's get started, and I'll turn things back over to my colleague Will Doyle.

*Will:*

Thanks Jeff. The webinar consists of a series of slides that we're going to talk through.

First, we will provide an overview of the WWC, the unique protocols that we developed for postsecondary research, and the special considerations for postsecondary topic areas.

Next, we will go into the specifics of some key features of designs eligible under WWC standards, so that is randomized control trials and some types of quasi-experimental designs.

After that, we will discuss best practices for reporting statistics. We'll also mention some habits or practices in reporting statistics that often prevent studies from meeting standards.

Then, we will summarize the key points, along with providing a set of resources you can use and ways you can stay connected to the WWC.

Finally, we will turn to your questions.

*Will:*

OK, let's start by briefly describing who we are, what we do and don't do, and why we exist. And by "we", I mean IES, Development Services Group, and our other partners who make up the WWC as a whole.

The Institute of Education Sciences was constituted to begin making investments in education research as a result of the No Child Left Behind Act. Among the initial investments of IES, the WWC was established in 2002 to be a central and trusted source of scientific evidence for what works in education. The Clearinghouse aims to identify all relevant research studies on an intervention, review those studies against design standards, and then synthesize findings from the highest quality research.

With this in mind, it is also important to clarify what the WWC does not do.

We don't directly test or study interventions; instead, we review existing research evaluating the effectiveness of interventions.

We don't commission new research on interventions.

And we don't recommend or approve specific interventions.

Instead, our goal is to let decision makers use the information we've generated to decide for themselves what they should do, given their needs and context.

*Will:*

Before we move on to how the WWC reviews studies, I want to take a brief moment to discuss some of the WWC products available in the Postsecondary Area.

On the slide, you will see the webpage that list all of the postsecondary products. Currently, the database houses 37. These range from intervention reports, to single study and quick reviews, and these go across a wide variety of topic areas and outcomes.

By typing in a keyword into the search bar, you are able to find products that pertain to certain outcomes or interventions.

*Will:*

I also want to mention the Reviewed Studies Database.

This database includes every study the WWC has reviewed and the rating the study has received. By searching for topics or authors in the search bar, you may find if the study has been rated and what rating it has been assigned.

The example listed on the screenshot is for “college enrollment”. By searching for these keywords, we can see that 16 different studies have been reviewed.

This is a great resource for grant competitions or if you just want to know how the WWC rates individual studies.

And with that, I will turn it over now to Jeff who will discuss the WWC’s reviewing standards and rating system. Jeff?

*Jeff:*

Thanks Will.

A large portion of our focus is on the assessment of individual studies against a set of design and analysis standards. These standards are explained in the WWC’s Handbook, and we’ll have a lot more to say on them today. But basically, the standards focus on the extent to which a study’s design and analysis can support causal claims. Therefore when a study meets WWC standards, it means that we are confident that any improvement in student outcomes is due to the intervention that was implemented and not to factors such as pre-existing differences between the intervention and comparison students.

To apply the standards, we train and certify reviewers, and only certified reviewers apply the standards for the WWC . Not all studies are eligible for review by the WWC, but among those that are, three ratings are possible:

The highest rating, *Meets WWC Group Design Standards Without Reservations*, is reserved for studies that provide the highest degree of confidence that an observed effect was caused by an intervention. Only RCTs with low levels of attrition can receive this rating.

The next rating, *Meets WWC Group Design Standards With Reservations*, is given to studies that provide a lower level of confidence that the observed effect was caused by the intervention.

The final rating, *Does Not Meet WWC Group Design Standards*, is given to studies that have not demonstrated a rigorous causal link between the intervention and the findings.

*Jeff:*

This infographic, which is available on the WWC’s website, illustrates how the standards are applied and how they eventually determine the rating of a research study. It also serves as a basis for the structure of this presentation.

Notice that there are four important questions that we ask:

At the top, the first question asks about the assignment of the groups – Is the assignment at random?

The second question asks about attrition, in other words, did anyone drop out of the study?

Third, the standards ask about the similarity of the intervention and comparison groups before the intervention started. This question only applies to QEDs and to high attrition RCTs.

The final question asks about confounds, or in other words, if there is anything other than the intervention that could account for the impact of the intervention.

These questions will all be discussed in detail during this webinar.

*Jeff:*

I mentioned earlier that the WWC has developed a Handbook that outlines the rules that the WWC follows when carrying out reviews; these rules include the standards against which intervention studies are evaluated. The Handbook was developed by IES in consultation with experts and went through a rigorous review process. It has been revised twice, and is now in its third version.

However, some aspects of the WWC's standards depend on the context of the research question, and therefore cannot be specified in the Handbook. To address this, the WWC develops documents known as "protocols". These protocols provide more specific guidance on the parameters of the review, such as sample eligibility characteristics, and on other elements related to the evaluation of studies.

Both the Handbook, and the associated protocols are available on the WWC's website.

Will?

*Will:*

Thanks for that overview Jeff. I think a little more context on protocols is needed here.

Currently the WWC has three protocols directly related to postsecondary education: the first is the "Review Protocol For Studies of Interventions for Developmental Students In Postsecondary Education" – which we have hyperlinked here, "Review Protocol for Studies of Interventions to Support the Transition to College", and "Review Protocol for Individual Studies in the Postsecondary Education Topic Area".

So, this last protocol is used for studies that are on topics that don't have their own protocol. For example, if the WWC conducted a review of a study of an intervention aimed at increasing critical thinking skills in the general college population, we would use this protocol to guide the review.

All of the postsecondary protocols have a standard set of outcomes that you see listed here. These all measure the extent to which students get into, make progress through, and successfully exit postsecondary education.

But some topic areas will have additional eligible outcomes. For example, when we review studies of interventions for students in developmental education, we are interested in the extent to which students make progress through developmental education.

In addition, protocols will also specify the covariates on which high attrition randomized experiments and all quasi-experiments the WWC will assess baseline comparability. You can see here that all of the postsecondary protocols require a measure of pre-intervention academic achievement and a pre-intervention measure of socioeconomic status for high attrition RCTs and all QEDs.

*Will:*

Thanks for that overview Jeff. I think a little more context on protocols is needed here.

Currently the WWC has three protocols directly related to postsecondary education: the first is the “Review Protocol For Studies of Interventions for Developmental Students In Postsecondary Education” – which we have hyperlinked here, “Review Protocol for Studies of Interventions to Support the Transition to College”, and “Review Protocol for Individual Studies in the Postsecondary Education Topic Area”.

So, this last protocol is used for studies that are on topics that don’t have their own protocol. For example, if the WWC conducted a review of a study of an intervention aimed at increasing critical thinking skills in the general college population, we would use this protocol to guide the review.

All of the postsecondary protocols have a standard set of outcomes that you see listed here. These all measure the extent to which students get into, make progress through, and successfully exit postsecondary education.

But some topic areas will have additional eligible outcomes. For example, when we review studies of interventions for students in developmental education, we are interested in the extent to which students make progress through developmental education.

In addition, protocols will also specify the covariates on which high attrition randomized experiments and all quasi-experiments the WWC will assess baseline comparability. You can see here that all of the postsecondary protocols require a measure of pre-intervention academic achievement and a pre-intervention measure of socioeconomic status for high attrition RCTs and all QEDs.

*Will:*

With the standards in mind, let’s turn to the specific criteria that make up the standards. As mentioned before, there are two general designs the WWC is interested in: Randomized Controlled Trials and Quasi-Experimental Designs.

Randomized controlled trials, or RCTs, provide the highest level of certainty about the effects of an intervention. The reason for this is that when participants are randomly assigned to the treatment or comparison group, we expect them to be similar on both measured and unmeasured characteristics.

Because we have a strong expectation that the only difference between the groups is exposure to the intervention, any difference in outcomes is attributable to the intervention. Studies that use an RCT design provide the greatest confidence in the results and this type of study can receive the highest rating, *Meets Group Design Standards Without Reservations*.

One thing that could affect the rating of an RCT, is the attrition, or dropout rate, of the individuals or units, such as schools, in the study. RCTs with high attrition can only receive a rating of *Meets Group Design Standards With Reservations*.

Now, in quasi-experimental designs, just like in an RCT, some participants receive the treatment and some do not, but the groups were not formed randomly. Instead, they were formed some other way. For example, they may have chosen their groups, or they may have been placed in their groups by an administrative procedure.

Because the participants are not randomly assigned to intervention or comparison groups, we can, at best, only show that the groups are similar on measured characteristics, such as prior academic achievement – like student high school GPA - or socio-economic status.

At the same time, there is no way we can show that the groups are similar on unmeasured characteristics. This means the groups might differ due to something other than just the intervention, and that difference might affect the results of the study.

As a result, these studies cannot receive the highest rating, but a high-quality QED can receive a rating of *Meets Group Design Standards With Reservations*.

*Will:*

So thinking about RCTs specifically, attrition is the greatest threat to the internal validity of an RCT. Attrition can affect the expectation that groups are comparable at the start of the intervention, and hence can affect the outcomes that are observed. The WWC assesses attrition in two ways. First, we examine the proportion of the overall sample that did not contribute to the measurement of the study's effects. This is often referred to as "overall attrition".

We also examine what is known as "differential attrition", which refers to the extent to which attrition rates differed across the study groups.

If both overall and differential attrition are at acceptable levels, then an RCT can achieve the WWC's highest rating. If both are not at acceptable levels, then an RCT could potentially be rated as *Meets WWC Group Design Standards with Reservations*, or *Does Not Meet WWC Group*

*Design Standards.* You can find more information about attrition and the WWC's cutoffs in the Handbook.

And with that I'll hand it over to Jeff.

*Jeff:*

Thanks Will. You can tell that attrition is an important consideration because it determines whether an RCT can receive the highest rating or not. However, other dimensions of study design and analysis are also important. We would like to illustrate these other dimensions using a quasi-experimental study as an example.

The context for this example is a Summer Bridge program and how well it helps students place into college-level coursework.

First, there are two groups – a group of pre-college students on the left who participated in a Summer Bridge program, and a group of pre-college students on the right who did not.

Second, there was a pretest, such as the ACCUPLACER or COMPASS, given to both groups at the same time, prior to the intervention during the summer of 2015. The study authors also collected information on student socio-economic status via the FAFSA Expected Family Contribution. We use this information to assess the extent of group similarity prior to the start of the study.

Third, there are no obvious confounding factors, and therefore the only systematic difference in what happened to the two groups is participation in the Summer Bridge program. We will discuss these confounding factors in more detail in a minute.

Finally, the study uses an eligible outcome to assess the effects of the intervention. In this case, the authors measured student scores on the placement exam when it was retaken, just before the start of the Fall 2015 semester.

Let's go over these four factors in more detail now.

*Jeff:*

So - the first factor we'll discuss is the need for distinct groups. For a study to be eligible for review by the WWC, it needs to have at least two groups: one group that receives the intervention and another that serves as a comparison group. Of course the rationale here is that the comparison group allows us to estimate what would have happened without the intervention.

Designs that compare an intervention group to itself at an earlier point in time are not eligible for review by the WWC. For example, a study might examine student scores on a reading inventory at the beginning of the semester and at the end of the semester. The WWC does not review this type of study. The reason is that without a comparison group that didn't receive the intervention, it is difficult to tell how much of any change in reading scores is attributable to the intervention.



It is important to emphasize that intervention and comparison groups can be created randomly or non-randomly. So in other words, a researcher does not have to use a randomized design to meet WWC standards.

Finally, we should mention that the sample can be a random sample or a simple convenience sample. And, groups can be created retrospectively -- for example, by using an administrative dataset with propensity score matching -- or prospectively, such as when a group or a person actively volunteers to be a part of an intervention.

*Jeff:*

Moving on, we now turn to the second factor: the similarity of the groups before the start of the intervention.

This has come up a lot lately because it is a critical part of evaluating studies that do not use randomization to place students into groups. The key reason for this is that we want to be sure the two groups are as similar as possible at the start of the intervention.

Imagine implementing a program that aims to improve the reading ability of students placed into developmental writing courses. At the end of the semester, the students in the intervention group outscore the students in the comparison group on a reading test. We would probably be tempted to conclude that the intervention was effective.

But imagine that it turns out that prior to the start of the intervention, the students in the intervention group were better readers, on average, than the students in the comparison group. We might have made a mistake when we concluded that the intervention was effective, and we would have had no idea that we might be wrong if the authors hadn't given the pretest.

As a result, the WWC requires that QEDs and high attrition RCTs demonstrate that the groups were reasonably similar prior to the intervention on key characteristics that are likely to be correlated with the outcomes. The specific characteristics that must be equated and the rules for demonstrating equivalence vary by review topic and are documented in the review protocol.

In most cases across the WWC as a whole, the WWC looks at equivalence on a *direct* pretest measure of the outcome. For example, measuring students' scores on a reading inventory before and after the intervention.

To assess pre-intervention comparability, the WWC looks at the size of the difference between the groups in standardized mean difference effect size units. The Handbook discusses how these effect sizes are calculated, but if you are familiar with Cohen's *d*, then the WWC's effect size will be familiar to you.

The chart on the bottom of the slide explains how this works and how group differences are handled.

As the column on the left indicates, groups that are no more than .05 standard deviations apart are considered to be equivalent. We can be confident that pre-existing differences on the observed measure will not overly influence the estimate of the intervention's impact.

The middle column indicates what happens when the groups have moderately sized differences, which we define as a difference between .05 and less than or equal to .25 standard deviations on the pretest measure. In this case, a study may meet standards **if and only if** the study authors control for these differences using an approved statistical technique. The approved techniques are described in the Handbook.

The column on the right indicates what happens when differences are greater than 0.25 standard deviations. In these cases, the WWC believes that the groups are too far apart on the pretest measure for the outcome to be meaningfully interpreted, and therefore the study does not meet WWC standards.

Finally, another reason a study might not meet standards is that we simply do not have enough information to assess how similar the groups were prior to the start of the intervention.

Not having this information is a common reason studies don't meet standards, and this seems particularly true for older studies. The WWC will reach out to the authors for this information if we think it might be available; however, we don't always receive it.

We strongly encourage you to review the WWC Reporting Guide for Study Authors for guidance on writing up your results; this guidance is hyperlinked on the "Resources" slide below. In addition, if you are interested in maximizing your chances for meeting WWC standards we also recommend you assess and report information about baseline equivalence, and adjust for any pre-existing differences in your estimation models. Let me emphasize again that the WWC's Handbook has a lot more information on these points.

Will?

*Will:*

Thanks, Jeff. To elaborate on the point about postsecondary education examples, let's discuss briefly how the WWC handles outcomes that do not have a direct pretest.

In my experience, postsecondary research studies are rarely as simple as the previous example. In cases where there is no pretest for an outcome, such as enrollment in a postsecondary institution, the WWC's Handbook requires that the protocol specify the pre-intervention variable or variables that need to be examined for comparability. All of the postsecondary protocols currently require that the WWC be able to assess pre-intervention comparability on a measure of prior academic achievement AND a measure of socioeconomic status. We chose these variables because we believe that they are the variables that are most highly correlated with the outcomes we are interested in.

*Will:*

Turning now from baseline equivalence, let's move to the third factor we want to assess: whether a component or factor of the study, other than the intervention, may influence the outcome, or in other words, if the study suffers from confounds. The easiest way to understand confounds is to talk through some of them.

The most common confound we see is that there is only one unit in one or both conditions. For instance, a new behavioral intervention is implemented in one college and outcomes are compared to the other colleges nearby.

Unfortunately, in this case we cannot isolate the effect of the intervention because the intervention and comparison colleges may differ in ways related to behavioral outcomes. One college may have a greater need for financial aid, different instructor qualifications, or different students attending, which could influence behavioral outcomes. To decrease the likelihood that there is some factor confounding the results, each of the intervention and comparison groups must contain at least two units.

Another confounding factor we often see is that there is one person providing the intervention who only interacts with the intervention group. For instance a single math instructor teaches her classes the same way and the comparison group are the other math instructors. In this case, the math instructor who is teaching the intervention group may be an incredibly dynamic teacher and her personality may affect student outcomes as much as the particular curriculum or strategy she is implementing. This study cannot meet standards as a test of the curriculum or strategy because we cannot separate how much of the observed effect was due to the intervention and how much was due to the particular math instructor.

Finally, another frequent confound is history. A history confound happens when pre-tests and post-tests are not assessed at the same time for both groups, such as when one year's students are compared to another year's students. When pretests and posttests are not assessed at the same time for both groups, there could be something occurring at same time as intervention, such as a change in curriculum or policies, which could be causing the observed outcomes. The point here is that in order for a study to meet standards, data have to be collected at the same time for both groups.

*Will:*

OK let's move now to the final of the four aspects, outcome measures. Outcomes are used to estimate the impact of the intervention, and as such, the WWC wants to ensure that the outcome measures meet some basic criteria.

First, an outcome must have face validity – that is, it appears to measure what it says it measures. Almost all outcomes meet this criterion.

Second, the scores produced by the measure must be reliable. Reliability ensures that the measure is consistent across time and people. The specific reliability requirements vary across

protocols and measures, so for more specific information here, consult one of the protocols we mentioned.

The third criteria for an outcome is that it not be overaligned with the intervention or comparison group. Overalignment occurs when outcome measures are more closely aligned with one of the research groups, intervention or comparison, than the other. For example, when outcome measures are closely aligned with or tailored to the intervention group, the intervention group may have an advantage over the comparison group simply from seeing the material on a previous occasion.

To help you think about this, assume we are looking at a study that aims to teach students to be better writers. The intervention involves giving students specific prompts to which they respond by writing a paragraph. Now, assume that the outcome measure involves giving both intervention and comparison students these same prompts, and judges evaluate the quality of the responses to the prompts. Because the intervention students had experience with the prompts, they have an advantage over the comparison students, and it would not be a surprise to see that they did better on the writing assignment. So, the WWC keeps an eye out for situations like this, and outcomes that are overaligned to the intervention receive a rating of *Does Not Meet WWC Group Design Standards*.

And with that, I'll turn it back over to Chris for a final overview of these criteria.

*Chris:*

Thanks, Will. Before we move on, I'll reinforce the takeaways to keep in mind about the WWC's reviewing criteria.

Start with at least two distinct groups, with data collected before and after the intervention.

Report attrition data including the number of participants in each group at each stage.

Assess baseline equivalence before estimating impacts and adjust for differences in baseline measures if necessary.

Avoid confounding factors such as single units or single implementers.

Use valid measures, make sure that the scores that are produced are reliable, and use outcomes that are not overaligned with the intervention.

Clearly document design, data collection, and analytic procedures.

Use WWC resources to help with measure selection, analysis, and reporting.

And with that, I'll turn it back over to Jeff to discuss some issues we see when studies report statistics.

Jeff?

*Jeff:*

Thanks Chris.

Well, as Chris mentioned, our final topic of the day is a short tutorial on what we're looking for in terms of the reported statistics.

We want to emphasize that a surprising portion of the studies we review fail to meet standards simply because inadequate statistics are reported. I'll also highlight some other things researchers can do to help us fully and accurately describe their studies.

On this slide, we have listed some of the common places where statistics are needed yet are often under-reported.

For example, it is very important in an RCT to report the sample size at randomization and at data collection.

For both randomized experiments and quasi-experiments, the baseline characteristics of the *analytic sample* are important. The analytic sample is the sample on which analyses are based. So, for instance, if 200 people are initially assigned to intervention and comparison groups, but only 180 complete the outcome measure, these 180 students are the analytic sample size.

In the calculation of effect sizes, the WWC uses sample sizes along with simple raw (or unadjusted) means and standard deviations for continuous measures. For binary measures, the WWC uses the relevant percentages in each group, for instance the percentage of Pell eligible students.

For the impact or outcome analyses, it is again important to report sample sizes and descriptive statistics for each group. These numbers are the way in which we calculate an effect size and therefore very important.

*Jeff:*

Let's turn to the reporting of sample sizes.

A participant flow chart, like the one presented on this slide, is a simple way to report the sample sizes at each important phase of the study. And although this information is specifically needed in the context of an RCT, it can be helpful for a QED as well.

You can see that the chart tracks sample sizes at three time points: eligibility assessment, assignment to conditions, and data analysis. The latter two time points are the most critical for the WWC - we use information about the assignment and analysis sample sizes to assess attrition in RCTs. As you'll see, we also use sample sizes at the time of assignment to help us assess

group similarity before the start of the intervention, and sample sizes at data analysis to help us compute effect sizes for the study's outcomes.

*Jeff:*

Moving on, as a follow-up to our discussion on baseline equivalence, let's turn to the reporting of baseline equivalence statistics. Again, baseline equivalence data are essential information for a QED in order to Meet Standards with Reservations.

As can be seen in the table, descriptive statistics are reported separately for the intervention and comparison groups. Here, the authors report high school GPA and FAFSA expected family contribution. The table also illustrates how a binary variable -- like whether or not students were assigned to developmental education -- can be displayed.

With this information, the WWC can now assess the similarity of the intervention and comparison groups before the study started.

While this may seem overly simple, we want to emphasize again that many studies we review fail to meet standards because these few pieces of information are missing or are incomplete.

Will?

*Will:*

Related to what Jeff just discussed on baseline equivalence, let's turn to the reporting of outcome analyses.

On the slide, we see that the authors have provided descriptive statistics for each group separately. In addition, the study authors provided the adjusted means, maybe from an ANCOVA or regression model, as well as the unadjusted means.

The key is that the authors provided the raw standard deviations. If the standard deviations reported derive from an ANCOVA model they may not be appropriate to use in the calculation of an effect size. We often have to request for the raw standard deviations from study authors to perform this calculation.

Finally, in the table, the authors provided the analytic sample sizes these statistics are based on.

From these descriptive statistics, we can fully report the effects on an intervention. If information is reported in this way, no additional information is needed from an author and the study will almost certainly be included in a relevant WWC product.

*Will:*

To give you an idea of the importance of the baseline equivalence issue, we gathered information on the postsecondary studies that the WWC has reviewed to date.

As you can see, of the approximately 200 studies we have reviewed so far, about 30% meet standards – this percentage is slightly higher than the WWC’s overall percentage of studies that meet standards (which is about 25%).

It’s interesting to note that by far, the most common reason studies do not meet standards is that there is a baseline equivalence problem. If you want your study to be used by the WWC, assess baseline equivalence prior to running analyses, and adjust for baseline differences if necessary -- and remember, in nonrandomized experiments, this is almost always necessary.

And with that I will hand it back over to Chris.

*Chris:*

Thank you Will. Very helpful information indeed.

To summarize, I’d like to conclude by reminding our audience where the Handbook, along with many other resources can be found.

Again, our *Handbook* provides a comprehensive description of all of our procedures and standards, so that anyone can understand how we work and how decisions are made.

The *Reporting Guide for Study Authors* provides some guidance on the types of information we need to be able to assess studies against standards.

The *Study Review Guide* is the tool we use when assessing studies. It is available to the public, as well as instructions for its use.

All of our *Review Protocols* are posted on the web, again to facilitate transparency.

And finally, if you’re looking for whether we’ve reviewed a study or what it was rated, we provide a *Database of Reviewed Studies* containing citations for and information on over 11,000 studies that have been identified for one of our reviews.

*Chris:*

You can also find more information about the What Works Clearinghouse by visiting the main website.

Front and center on the WWC’s homepage is a set of rotating images that highlight recent products or resources around a theme. These images are regularly updated. For example, we recently posted a What Works Clearinghouse Year in Review guide to 2015. It includes all the major product launches from the past year and is a great place to start to learn more about what we do.

As always, you can receive regular updates on new WWC products through social media by friending us on Facebook or following us on Twitter.

And you can also sign up to receive email newsflashes. Click on News & Events, and select Newsflash. Enter your email address and then expand NCEE, and check What Works Clearinghouse.

*Chris:*

Finally, you can also help us improve the Clearinghouse. If you have a suggestion or question, the best way to reach us is through our Help Desk. You can get to the Help Desk through our Contact Us page, which can be found under About Us.

With that, I'd like to thank our presenters today, Jeff Will, and Sandra, as well as the staff members at Development Services Group who helped put this together, for an excellent and informative talk.

To finalize, let's wrap up by taking some of your questions. For those of you who missed the beginning, if you want to submit questions, use the chat tool on your webinar software.

To help us get through the questions, I'll turn it over to Sandra who will talk through the questions we have received during the webinar. Sandra?