From: Zalman Usiskin [z-usiskin@uchicago.edu]
Sent: Thursday, July 21, 2011 3:18 PM
To: What Works
Cc: denisse@usf.edu
Subject: Re: WWC High School Math Review of University of Chicago School
Mathematics Project 6-12 Curriculum

Dear Ms. Constantine:

Thank you for this notification.

We appreciate the clarity and completeness of your overall report and we appreciate your acceptance of two studies as meeting WWC criteria.

We think, however, that you have mistakenly excluded two others. The reports of the evaluations of UCSMP Advanced Algebra (Thompson, Senk, Witonsky, Usiskin, & Kaeley, 2001) and UCSMP Geometry (Thompson, Witonsky, Senk, Usiskin, & Kaeley, 2003) each contain a detailed description of the strict procedures used to determine whether supposedly matched pairs of treatment and comparison classes were actually equivalent at the beginning of the year and to eliminate from the analysis those pairs of classes that were not. You might wish to look again at the reports of these studies.

Zalman Usiskin
Professor Emeritus of Education
Director, UCSMP
The University of Chicago
1225 E. 60th St.
Chicago, IL 60637

What Works wrote:
>
> Dear Dr. Usiskin and Dr. Thompson,
>
>
>
> The attached letter is to notify you that the What Works Clearinghouse
> (WWC) has completed the review of the research on /University of
> Chicago School Mathematics Project (UCSMP) 6-12 Curriculum/ and
> determined that this intervention is eligible for an intervention
> report according to the High School Math review protocol. We have
> also attached a courtesy copy of the report which will be posted on
> the WWC website on July 19, 2011. As a reminder, this report is
> covered by the embargo agreement sent to you on May 12, 2010,
> requiring you not to copy, distribute, or discuss the report with
> members of the public outside your organization, prior to release of
> the report by the Institute of Education Sciences.
>
>
>
> Sincerely,

>
>
>
> Jill Constantine
>
> Director, What Works Clearinghouse
>
>
>

From: WhatWorks
Sent: Friday, July 22, 2011 3:18 PM
To: 'Zalman Usiskin'
Subject: RE: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Dr. Usiskin,

Thank you for contacting the What Works Clearinghouse (WWC). We have received your email below. WWC staff are reviewing your request and will prepare a response.

What Works Clearinghouse

The What Works Clearinghouse was established by the U.S. Department of Education's Institute of Education Sciences to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education. For more information, please visit http://ies.ed.gov/ncee/wwc/.

-----Original Message-----
From: Zalman Usiskin [mailto:z-usiskin@uchicago.edu]
Sent: Thursday, July 21, 2011 2:18 PM
To: What Works
Cc: denisse@usf.edu
Subject: Re: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Ms. Constantine:

Thank you for this notification.

We appreciate the clarity and completeness of your overall report and we appreciate your acceptance of two studies as meeting WWC criteria.

We think, however, that you have mistakenly excluded two others. The reports of the evaluations of UCSMP Advanced Algebra (Thompson, Senk, Witonsky, Usiskin, & Kaeley, 2001) and UCSMP Geometry (Thompson, Witonsky, Senk, Usiskin, & Kaeley, 2003) each contain a detailed description of the strict procedures used to determine whether supposedly matched pairs of treatment and comparison classes were actually equivalent at the beginning of the year and to eliminate from the analysis those pairs of classes that were not. You might wish to look again at the reports of these studies.

Zalman Usiskin
Professor Emeritus of Education
Director, UCSMP
The University of Chicago
1225 E. 60th St.
Chicago, IL 60637

What Works wrote:
>
> Dear Dr. Usiskin and Dr. Thompson,
>
>
>
> The attached letter is to notify you that the What Works Clearinghouse
> (WWC) has completed the review of the research on /University of
> Chicago School Mathematics Project (UCSMP) 6-12 Curriculum/ and
> determined that this intervention is eligible for an intervention
> report according to the High School Math review protocol. We have
> also attached a courtesy copy of the report which will be posted on
> the WWC website on July 19, 2011. As a reminder, this report is
> covered by the embargo agreement sent to you on May 12, 2010,
> requiring you not to copy, distribute, or discuss the report with
> members of the public outside your organization, prior to release of
> the report by the Institute of Education Sciences.
>
>
>
> Sincerely,
>
>
>
> Jill Constantine
>
> Director, What Works Clearinghouse
>
>
>

From: WhatWorks
Sent: Wednesday, August 10, 2011 4:20 PM
To: 'Zalman Usiskin'
Subject: RE: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Dr. Usiskin,

The WWC Quality Review Team is reviewing your email and will prepare a written response. The Quality Review Team responds to concerns raised by study authors, curriculum developers or other relevant parties about WWC reviews published on our website. These quality reviews are undertaken when concerned parties present evidence that a WWC review may be inaccurate. When a quality review is conducted, a researcher who was not involved in the initial review undertakes an independent assessment of the study in question. The researcher also investigates the procedures used and decisions made during the original review of the study. If a quality review concludes that the original review contained errors, a revision will be published. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the What Works Clearinghouse.

Thank you,

What Works Clearinghouse

From: WhatWorks
Sent: Friday, July 22, 2011 2:18 PM
To: 'Zalman Usiskin'
Subject: RE: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Dr. Usiskin,

Thank you for contacting the What Works Clearinghouse (WWC). We have received your email below. WWC staff are reviewing your request and will prepare a response.

What Works Clearinghouse

-----Original Message-----
From: Zalman Usiskin [mailto:z-usiskin@uchicago.edu]

Sent: Thursday, July 21, 2011 2:18 PM
To: What Works
Cc: denisse@usf.edu
Subject: Re: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Ms. Constantine:

Thank you for this notification.

We appreciate the clarity and completeness of your overall report and we appreciate your acceptance of two studies as meeting WWC criteria.

We think, however, that you have mistakenly excluded two others. The reports of the evaluations of UCSMP Advanced Algebra (Thompson, Senk, Witonsky, Usiskin, & Kaeley, 2001) and UCSMP Geometry (Thompson, Witonsky, Senk, Usiskin, & Kaeley, 2003) each contain a detailed description of the strict procedures used to determine whether supposedly matched pairs of treatment and comparison classes were actually equivalent at the beginning of the year and to eliminate from the analysis those pairs of classes that were not. You might wish to look again at the reports of these studies.

Zalman Usiskin
Professor Emeritus of Education
Director, UCSMP
The University of Chicago
1225 E. 60th St.
Chicago, IL 60637

What Works wrote:
>
> Dear Dr. Usiskin and Dr. Thompson,
>
>
>
> The attached letter is to notify you that the What Works Clearinghouse
> (WWC) has completed the review of the research on /University of
> Chicago School Mathematics Project (UCSMP) 6-12 Curriculum/ and
> determined that this intervention is eligible for an intervention
> report according to the High School Math review protocol. We have
> also attached a courtesy copy of the report which will be posted on
> the WWC website on July 19, 2011. As a reminder, this report is
> covered by the embargo agreement sent to you on May 12, 2010,
> requiring you not to copy, distribute, or discuss the report with
> members of the public outside your organization, prior to release of
> the report by the Institute of Education Sciences.
>
>
>

> Sincerely,
>
>
>
> Jill Constantine
>
> Director, What Works Clearinghouse
>
>
>

From: Zalman Usiskin <z-usiskin@uchicago.edu>
Sent: Wednesday, October 05, 2011 7:44 PM
To: What Works
Cc: Denisse Thompson
Subject: Re: What Works Clearinghouse (WWCPC 2011002)

Dear Ms. Constantine:

Thank you very much for your detailed letter of September 29th. We very much appreciate the attention you gave to the issue I raised. We did not compute effect sizes and were unaware that you were using this criterion to accept or reject a study based on its design. Statistical controls still may not account for differences in expectations that two different teachers might have for their students, so even a very strict criterion is not a perfect sorter of studies, but we do understand what you did.

We are now computing the effect sizes for the pretests in the third edition studies that we have undertaken. Since these studies have already been completed, it will be interesting to see how many satisfy your criterion.

Zalman Usiskin
Professor Emeritus of Education
Director, University of Chicago School Mathematics Project
The University of Chicago
1225 E. 50th St.
Chicago, IL 60637

What Works wrote:
Dear Dr. Usiskin,

Attached is a response to the questions you raised in your July 21 message to the What Works Clearinghouse (WWC).

Thank you,

What Works Clearinghouse

From: WhatWorks
Sent: Wednesday, August 10, 2011 4:20 PM
To: 'Zalman Usiskin'
Subject: RE: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Dr. Usiskin,

The WWC Quality Review Team is reviewing your email and will prepare a written response. The Quality Review Team responds to concerns raised by study authors, curriculum developers or other relevant parties about WWC reviews published on our website. These quality reviews are undertaken when concerned parties present evidence that a WWC review may be inaccurate. When a quality review is conducted, a researcher who was not involved in the initial review undertakes an independent assessment of the study in question. The researcher also investigates the procedures used and decisions made during the original review of the study. If a quality review concludes that the original review contained errors, a revision will be published. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the What Works Clearinghouse.

Thank you,

What Works Clearinghouse

From: WhatWorks
Sent: Friday, July 22, 2011 2:18 PM
To: 'Zalman Usiskin'
Subject: RE: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Dr. Usiskin,

Thank you for contacting the What Works Clearinghouse (WWC). We have received your email below. WWC staff are reviewing your request and will prepare a response.

What Works Clearinghouse

-----Original Message-----
From: Zalman Usiskin [mailto:z-usiskin@uchicago.edu]
Sent: Thursday, July 21, 2011 2:18 PM
To: What Works
Cc: denisse@usf.edu
Subject: Re: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Ms. Constantine:

Thank you for this notification.

We appreciate the clarity and completeness of your overall report and we appreciate your acceptance of two studies as meeting WWC criteria.

We think, however, that you have mistakenly excluded two others. The reports of the evaluations of UCSMP Advanced Algebra (Thompson, Senk, Witonsky, Usiskin, & Kaeley, 2001) and UCSMP Geometry (Thompson, Witonsky, Senk, Usiskin, & Kaeley, 2003) each contain a detailed description of the strict procedures used to determine whether supposedly matched pairs of treatment and comparison classes were actually equivalent at the beginning of the year and to eliminate from the analysis those pairs of classes that were not. You might wish to look again at the reports of these studies.

Zalman Usiskin
Professor Emeritus of Education
Director, UCSMP
The University of Chicago
1225 E. 60th St.
Chicago, IL 60637

What Works wrote:
>
> Dear Dr. Usiskin and Dr. Thompson,
>
>
>
> The attached letter is to notify you that the What Works Clearinghouse
> (WWC) has completed the review of the research on /University of
> Chicago School Mathematics Project (UCSMP) 6-12 Curriculum/ and
> determined that this intervention is eligible for an intervention
> report according to the High School Math review protocol. We have
> also attached a courtesy copy of the report which will be posted on
> the WWC website on July 19, 2011. As a reminder, this report is
> covered by the embargo agreement sent to you on May 12, 2010,
> requiring you not to copy, distribute, or discuss the report with
> members of the public outside your organization, prior to release of
> the report by the Institute of Education Sciences.
>
>
>
> Sincerely,
>
>
>
> Jill Constantine
>
> Director, What Works Clearinghouse
>

>
>

From: Sakari Morvey
Sent: Wednesday, June 23, 2010 3:46 PM
To: What Works
Cc: 'Loyco, Katherine B'
Subject: FW: WWC High School Math Review of University of Chicago School Math Project

Leah,

Can you create a new issue for this?

From: Jean Knab
Sent: Wed 6/23/2010 3:45 PM
To: 'WhatWorks'
Cc: Sakari Morvey
Subject: FW: WWC High School Math Review of University of Chicago School Math Project

Here is the letter.

-----Original Message-----
From: Neil Seftor
Sent: Wednesday, June 16, 2010 9:44 AM
To: Scott Cody; Jean Knab
Subject: FW: WWC High School Math Review of University of Chicago School Math Project

This is the response we received to a developer letter for HSM. It's full of all sorts of criticisms of the WWC. Is this something you think needs an additional response?

-----Original Message-----
From: Claire Smither
Sent: Wednesday, June 16, 2010 9:07 AM
To: Neil Seftor; Christina Tuttle
Subject: FW: WWC High School Math Review of University of Chicago School Math Project

We have received a response to the UCSMP Developer Letter ...

-----Original Message-----
From: Zalman Usiskin [mailto:z-usiskin@uchicago.edu]
Sent: Tuesday, June 15, 2010 11:02 PM
To: What Works
Cc: Denisse Thompson
Subject: Re: WWC High School Math Review of University of Chicago School Math Project

Dear Dr. Seftor:

We very much appreciate the desire of the WWC to review UCSMP materials.

My delay in responding to your earlier e-mail was due to our experiences with previous reviews of studies of UCSMP materials at the middle and high school levels. In my opinion, these reviews have been conducted with a naiveté that boggles the mind of someone who has been doing comparative studies of curriculum for over 40 years. We have crafted what we hope is a response that applies not only to UCSMP materials but to all of the materials you consider.

Here are five of the major weaknesses.

(1) There is little if any conception of what constitutes a good sample for a typical curriculum study. Specifically, the WWC reviews do not seem to understand that the appropriate bite size for a curriculum study is not the student, but the classroom, for it is typically the classroom that is given a treatment. As a result, the reviewers look upon studies with one classroom (n = 1) and somehow think that this gives a reliable picture of how materials will fare in others. The naiveté here is that anyone who has implemented curriculum in multiple classrooms and in multiple schools knows that the same curriculum can fare well in one place and not so well in another. Teachers and school climate influence the implemented and learned curriculum in a wide variety of ways.

The result of this first weakness is that the WWC reviews include studies that are too small to be generalizable in any way and exclude large-scale studies that by their very size do not present the exact same conditions in the implementation of the materials but more accurately reflect the variety of ways in which materials are used and are more likely to have generalizable results.

(2) There is no consideration of the treatment in the "control" group. There does not seem even to be any realization of what would seem to be obvious, that in a comparison study the comparative performance of two treatments depends on the quality of both treatments, not just the quality of the "new" treatment being examined.

Thus, for instance, if one is testing new materials in a school where the performance is historically low, there is more room for the new materials to perform better than the old than if the testing takes place in a school where the performance is historically high. And, in fact, since there are a great number of curricula in use, if there were some order to the quality of curricula, it is possible for a study to decide that curriculum X was not as effective as curriculum Y and not realize that curriculum X might be better than every other available curriculum.

Thus, to pass judgment on a curriculum on the basis of its performance against only one other curriculum can at best tell only a small part of the story of the relative effectiveness of the curriculum. If curriculum X is less effective than curriculum Y, X could be anywhere from the worst to the second best curriculum available. If it is more effective, it could be anywhere from the best to the second worst available.

(3) The reviewers do not seem to be able to distinguish studies that are formative in nature - i.e., designed to inform the curriculum developers of how their materials are faring so that the materials can be improved

before publication - and those that are summative in nature - i.e., with published materials. That is, the conditions under which studies are undertaken are usually not given any consideration.

The result of this third weakness is that the WWC reviews often consider studies of test and pilot editions of materials and ignore the existence of later versions of the same materials that may have taken into account the results of the very studies that are being reviewed.

(4) With respect to instruments, the reviewers seem to assume that all curricula are trying to cover the same content and that the differences are mainly instructional. With projects like ours, this is an invalid assumption - indeed, we would not exist if all we are trying to do is teach the old curriculum better. We are attempting to modernize the curriculum, to bring it up to date, to include new ideas and exclude old ones. The experiences our students have are quite different from the experiences of students in the comparable existing courses. When students have not had an opportunity to learn content (whether old or new), one has to be very careful about giving them a test over that content, because the student may simply give up and not even attempt the items that the student feels have not been covered. Furthermore, some of the content is so new that there are no standardized tests that cover it. For these reasons, we have to very carefully analyze test results with respect to opportunity to learn and we have to devise some of our own tests.

The result of this fourth weakness is that the WWC reviews of new materials often do not take into account how students do on the content that is new in these materials, that is, on the content that makes them the most different.

(5) An additional issue to consider when studying high school treatments is that beyond algebra I, there have not been current standardized tests that are content specific. That is, there have not been standardized tests focusing on geometry or algebra II or courses beyond algebra II. (We understand that Achieve is currently working to develop such a test at the algebra II level.) Thus, projects often have no opportunity to use a standardized instrument at these levels unless it is a general test, such as the SAT or the ITED-Q. At UCSMP, although we have developed our own tests, we have the classroom teachers indicate whether they taught or reviewed the content needed for students to answer the items and analyze the results based on teachers' responses, both the UCSMP teachers and comparison teachers. We believe this is a transparent way to reduce bias. Yet, WWC completely discounts such studies. We believe this is unfortunate, particularly because we have data to indicate that there have often been many items on a standardized test, even at algebra I, for which teachers in both UCSMP and comparison treatments did not teach the necessary content.

I hope the WWC in the Obama administration is correcting these and other flaws that permeated WWC reviews in the Bush administration. With that in mind, we are responding to your specific requests.

LIST OF STUDIES:
On the existing list, we have the following comments:

• McConnell, K. This reference should be McConnell, J.
• School reform math programs. We have no idea what this is and the reference is too incomplete to lead us anywhere.
• Secondary options and post-secondary expectations. We also have no idea what this is.
• Swann. If you have included this study, why did you not include the far larger studies done by UCSMP (see below)?
• Thompson, D.R., et al. (1st listing): There is a typo in the second line.
• Thompson, D.R., et al. (3rd listing): This duplicates the previous reference. Perhaps you meant the 2nd edition study of UCSMP Transition Mathematics (see below).

MISSING STUDIES:
UCSMP research during the development of its first edition:
Hedges, Larry V., Susan S. Stodolsky, Sandra Mathison, and Penelope V. Flores (1986). Transition Mathematics Field Study. Chicago: University of Chicago School Mathematics Project. Summarized in Transition Mathematics Teacher's Edition, Scott Foresman (1990), pp. T46-T50. Full report is available from UCSMP.

Hirschhorn, Daniel B., and Sharon Senk (1992). Calculators in the UCSMP Curriculum for Grades 7 and 8. In Calculators in Mathematics Education, edited by James T. Fey and Christian R. Hirsch. Reston, VA: NCTM.

Senk, Sharon L. (1989). Assessing Students' Knowledge of Functions. In Proceedings of the Eleventh Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, vol. 2, edited by C.A. Mahrer, G.A. Goldin, and R.B. Davis. New Brunswick, NJ: Rutgers University.

Thompson, Denisse Rubilee (1992). An Evaluation of a New Course in Precalculus and Discrete Mathematics. Volumes I and II. Doctoral dissertation, The University of Chicago. DAI-A 54/08: 2934 (February 1994).

Thompson, Denisse R. and Sharon L. Senk (1993). Assessing reasoning and proof in high school. In Assessment in the Mathematics Classroom, edited by Norman L. Webb and Arthur F. Coxford, 1993 Yearbook of the National Council of Teachers of Mathematics, pp. 167-176. Reston, VA: NCTM.

UCSMP second edition evaluations:
Thompson, Denisse R., David Witonsky, Sharon L. Senk, Zalman Usiskin, Gurcharn Kaeley (2003). Evaluation of the Second Edition of UCSMP Geometry. Chicago: The University of Chicago School Mathematics Project. (Note: This report is available from UCSMP.)

Senk, Sharon L. (2003). Effects of the UCSMP Secondary Curriculum on Students' Achievement. In Standards-based School Mathematics Curricula: What Are They? What Do Students Learn?, edited by Sharon L. Senk and Denisse R. Thompson. Mahway, NJ: Lawrence Erlbaum, pp. 425-456.

Senk, S. L., & Thompson, D. R. (2006). Brief Report: Strategies used by second-year algebra students to solve problems. Journal for Research in Mathematics Education, 37, 116-128.

Thompson, Denisse R. and Sharon L. Senk. The Effects of Curriculum on Achievement in Second-Year Algebra: The Example of the University of Chicago School Mathematics Project (2001). Journal for Research in Mathematics Education 32: 58-84.

Other studies
Zahrt, Lori Tubbergen (2001). High School Reform Math Programs: An Evaluation for Leaders. Doctoral dissertation, Eastern Michigan University, 2001. (This research compared UCSMP to Core-Plus. The UCSMP students were a higher-performing group to begin with, judging from junior high school scores on the MAEP. This could account for some of the extraordinary differences favoring UCSMP.)

INTERVENTION SUMMARY
Developer: Should be University of Chicago School Mathematics Project. (We are a part of the University of Chicago, but it seems best to be more specific.)

Contact:
Listed is my address but the e-mail for Denisse Thompson. It would be good for you to list both of us as contacts and complete addresses for both. Here is complete contact information:
Zalman Usiskin, Professor Emeritus of Education
Director, UCSMP
6030 South Ellis Avenue
Chicago, IL 60637
773 702-1560
z-usiskin@uchicago.edu.

Denisse Thompson, Professor of Mathematics Education
Department of Secondary Education
University of South Florida
4202 East Fowler Ave., #DU105
Tampa, FL 33620-5650
813 974-2687
denisse@usf.edu

Cost:
The amount you have posted for teacher materials is not accurate and reflects ignorance of ancillary materials and their purchase. Your total seems to have been found by adding the individual costs of all material available for schools to purchase, including both package prices for sets of materials and the costs of the individual materials themselves. The reason or buying the packages is so that one does not need to buy the materials individually. Furthermore, the packages are typically given free to schools that adopt the text materials.

I would be happy to try to clarify any unclear information in this e-mail. Thank you again for the opportunity to respond.

Zalman Usiskin

What Works wrote:
>
> Dear Mr. Usiskin,
>
> The What Works Clearinghouse (WWC), an initiative of the U. S.
> Department of Education's Institute of Education Sciences, was
> established to provide educators, policymakers, researchers, and the
> public with a central and trusted source of scientific evidence of
> what works in education. As such, we review studies on education
> interventions that may be included in our reports.
>
> The purpose of the attached letter is to notify you that we are in the
> process of conducting a review of High School Math interventions and
> may be including University of Chicago School Math Project in our
> review. The High School Math review focuses on mathematics
> interventions for high school students in grades 9-12 designed to
> impact student achievement, including curriculum-based interventions,
> instructional techniques, and products designed to deliver content and
> monitor student progress. In this letter, we ask you to review a list
> of studies, review a brief intervention summary, and sign an embargo
> agreement.
>
> We originally emailed for your input on May 12^th , 2010. The original
> deadline for a response has passed. We are trying to maintain a brief
> pace, and ask that you respond as soon as possible.
>
> Sincerely,
>
> Neil Seftor, Ph.D.
>
> Principal Investigator, WWC High School Math review
>
> <<UCSMP Developer Letter.pdf>>
>

From: WhatWorks
Sent: Wednesday, June 23, 2010 3:04 PM
To: 'z-usiskin@uchicago.edu'
Subject: RE: WWC High School Math Review of University of Chicago School Math Project (WWCPC 2237)

Dear Dr. Usiskin,

Thank you for contacting the What Works Clearinghouse (WWC). We have received your email. WWC staff are reviewing your email and will prepare a response.

What Works Clearinghouse

The What Works Clearinghouse was established by the U.S. Department of Education's Institute of Education Sciences to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education. For more information, please visit http://ies.ed.gov/ncee/wwc/.

-----Original Message-----
From: Zalman Usiskin [mailto:z-usiskin@uchicago.edu]
Sent: Tuesday, June 15, 2010 11:02 PM
To: What Works
Cc: Denisse Thompson
Subject: Re: WWC High School Math Review of University of Chicago School Math Project

Dear Dr. Seftor:

We very much appreciate the desire of the WWC to review UCSMP materials.

My delay in responding to your earlier e-mail was due to our experiences with previous reviews of studies of UCSMP materials at the middle and high school levels. In my opinion, these reviews have been conducted with a naiveté that boggles the mind of someone who has been doing comparative studies of curriculum for over 40 years. We have crafted what we hope is a response that applies not only to UCSMP materials but to all of the materials you consider.

Here are five of the major weaknesses.

(1) There is little if any conception of what constitutes a good sample for a typical curriculum study. Specifically, the WWC reviews do not seem to understand that the appropriate bite size for a curriculum study is not the student, but the classroom, for it is typically the classroom that is given a treatment. As a result, the reviewers look upon studies with one classroom (n = 1) and somehow think that this gives a reliable picture of how materials will fare in others. The naiveté here is that anyone who has implemented curriculum in multiple classrooms and in multiple schools knows that the same curriculum can fare well in one place and not so well in another. Teachers and school climate influence the implemented and learned curriculum in a wide variety of ways.

The result of this first weakness is that the WWC reviews include studies that are too small to be generalizable in any way and exclude

large-scale studies that by their very size do not present the exact same conditions in the implementation of the materials but more accurately reflect the variety of ways in which materials are used and are more likely to have generalizable results.

(2) There is no consideration of the treatment in the " control" group. There does not seem even to be any realization of what would seem to be obvious, that in a comparison study the comparative performance of two treatments depends on the quality of both treatments, not just the quality of the " new" treatment being examined.

Thus, for instance, if one is testing new materials in a school where the performance is historically low, there is more room for the new materials to perform better than the old than if the testing takes place in a school where the performance is historically high. And, in fact, since there are a great number of curricula in use, if there were some order to the quality of curricula, it is possible for a study to decide that curriculum X was not as effective as curriculum Y and not realize that curriculum X might be better than every other available curriculum.

Thus, to pass judgment on a curriculum on the basis of its performance against only one other curriculum can at best tell only a small part of the story of the relative effectiveness of the curriculum. If curriculum X is less effective than curriculum Y, X could be anywhere from the worst to the second best curriculum available. If it is more effective, it could be anywhere from the best to the second worst available.

(3) The reviewers do not seem to be able to distinguish studies that are formative in nature - i.e., designed to inform the curriculum developers of how their materials are faring so that the materials can be improved before publication - and those that are summative in nature - i.e., with published materials. That is, the conditions under which studies are undertaken are usually not given any consideration.

The result of this third weakness is that the WWC reviews often consider studies of test and pilot editions of materials and ignore the existence of later versions of the same materials that may have taken into account the results of the very studies that are being reviewed.

(4) With respect to instruments, the reviewers seem to assume that all curricula are trying to cover the same content and that the differences are mainly instructional. With projects like ours, this is an invalid assumption - indeed, we would not exist if all we are trying to do is teach the old curriculum better. We are attempting to modernize the curriculum, to bring it up to date, to include new ideas and exclude old ones. The experiences our students have are quite different from the experiences of students in the comparable existing courses. When students have not had an opportunity to learn content (whether old or new), one has to be very careful about giving them a test over that content, because the student may simply give up and not even attempt the items that the student feels have not been covered. Furthermore, some of the content is so new that there are no standardized tests that cover it. For these reasons, we have to very carefully analyze test results with respect to opportunity to learn and we have to devise some of our own tests.

The result of this fourth weakness is that the WWC reviews of new materials often do not take into account how students do on the content that is new in these materials, that is, on the content that makes them the most different.

(5) An additional issue to consider when studying high school treatments is that beyond algebra I, there have not been current standardized tests that are content specific. That is, there have not been standardized tests focusing on geometry or algebra II or courses beyond algebra II. (We understand that Achieve is currently working to develop such a test at the algebra II level.) Thus, projects often have no opportunity to use a standardized instrument at these levels unless it is a general test, such as the SAT or the ITED-Q. At UCSMP, although we have developed our own tests, we have the classroom teachers indicate whether they taught or reviewed the content needed for students to answer the items and analyze the results based on teachers' responses, both the UCSMP teachers and comparison teachers. We believe this is a transparent way to reduce bias. Yet, WWC completely discounts such studies. We believe this is unfortunate, particularly because we have data to indicate that there have often been many items on a standardized test, even at algebra I, for which teachers in both UCSMP and comparison treatments did not teach the necessary content.

I hope the WWC in the Obama administration is correcting these and other flaws that permeated WWC reviews in the Bush administration. With that in mind, we are responding to your specific requests.

LIST OF STUDIES:
On the existing list, we have the following comments:

- McConnell, K. This reference should be McConnell, J.
- School reform math programs. We have no idea what this is and the reference is too incomplete to lead us anywhere.
- Secondary options and post-secondary expectations. We also have no idea what this is.
- Swann. If you have included this study, why did you not include the far larger studies done by UCSMP (see below)?
- Thompson, D.R., et al. (1st listing): There is a typo in the second line.
- Thompson, D.R., et al. (3rd listing): This duplicates the previous reference. Perhaps you meant the 2nd edition study of UCSMP Transition Mathematics (see below).

MISSING STUDIES:
UCSMP research during the development of its first edition:
Hedges, Larry V., Susan S. Stodolsky, Sandra Mathison, and Penelope V. Flores (1986). Transition Mathematics Field Study. Chicago: University of Chicago School Mathematics Project. Summarized in Transition Mathematics Teacher's Edition, Scott Foresman (1990), pp. T46-T50. Full report is available from UCSMP.

Hirschhorn, Daniel B., and Sharon Senk (1992). Calculators in the UCSMP Curriculum for Grades 7 and 8. In Calculators in Mathematics Education, edited by James T. Fey and Christian R. Hirsch. Reston, VA: NCTM.

Senk, Sharon L. (1989). Assessing Students' Knowledge of Functions. In Proceedings of the Eleventh Annual Meeting of the North American Chapter

of the International Group for the Psychology of Mathematics Education, vol. 2, edited by C.A. Mahrer, G.A. Goldin, and R.B. Davis. New Brunswick, NJ: Rutgers University.

Thompson, Denisse Rubilee (1992). An Evaluation of a New Course in Precalculus and Discrete Mathematics. Volumes I and II. Doctoral dissertation, The University of Chicago. DAI-A 54/08: 2934 (February 1994).

Thompson, Denisse R. and Sharon L. Senk (1993). Assessing reasoning and proof in high school. In Assessment in the Mathematics Classroom, edited by Norman L. Webb and Arthur F. Coxford, 1993 Yearbook of the National Council of Teachers of Mathematics, pp. 167-176. Reston, VA: NCTM.

UCSMP second edition evaluations:
Thompson, Denisse R., David Witonsky, Sharon L. Senk, Zalman Usiskin, Gurcharn Kaeley (2003). Evaluation of the Second Edition of UCSMP Geometry. Chicago: The University of Chicago School Mathematics Project. (Note: This report is available from UCSMP.)

Senk, Sharon L. (2003). Effects of the UCSMP Secondary Curriculum on Students' Achievement. In Standards-based School Mathematics Curricula: What Are They? What Do Students Learn?, edited by Sharon L. Senk and Denisse R. Thompson. Mahway, NJ: Lawrence Erlbaum, pp. 425-456.

Senk, S. L., & Thompson, D. R. (2006). Brief Report: Strategies used by second-year algebra students to solve problems. Journal for Research in Mathematics Education, 37, 116-128.

Thompson, Denisse R. and Sharon L. Senk. The Effects of Curriculum on Achievement in Second-Year Algebra: The Example of the University of Chicago School Mathematics Project (2001). Journal for Research in Mathematics Education 32: 58-84.

Other studies
Zahrt, Lori Tubbergen (2001). High School Reform Math Programs: An Evaluation for Leaders. Doctoral dissertation, Eastern Michigan University, 2001. (This research compared UCSMP to Core-Plus. The UCSMP students were a higher-performing group to begin with, judging from junior high school scores on the MAEP. This could account for some of the extraordinary differences favoring UCSMP.)

INTERVENTION SUMMARY
Developer: Should be University of Chicago School Mathematics Project. (We are a part of the University of Chicago, but it seems best to be more specific.)

Contact:
Listed is my address but the e-mail for Denisse Thompson. It would be good for you to list both of us as contacts and complete addresses for both. Here is complete contact information:
Zalman Usiskin, Professor Emeritus of Education
Director, UCSMP
6030 South Ellis Avenue
Chicago, IL 60637

773 702-1560
z-usiskin@uchicago.edu.

Denisse Thompson, Professor of Mathematics Education
Department of Secondary Education
University of South Florida
4202 East Fowler Ave., #DU105
Tampa, FL 33620-5650
813 974-2687
denisse@usf.edu

Cost:
The amount you have posted for teacher materials is not accurate and reflects ignorance of ancillary materials and their purchase. Your total seems to have been found by adding the individual costs of all material available for schools to purchase, including both package prices for sets of materials and the costs of the individual materials themselves. The reason or buying the packages is so that one does not need to buy the materials individually. Furthermore, the packages are typically given free to schools that adopt the text materials.

I would be happy to try to clarify any unclear information in this e-mail. Thank you again for the opportunity to respond.

Zalman Usiskin

What Works wrote:
>
> Dear Mr. Usiskin,
>
> The What Works Clearinghouse (WWC), an initiative of the U. S.
> Department of Education' s Institute of Education Sciences, was
> established to provide educators, policymakers, researchers, and the
> public with a central and trusted source of scientific evidence of
> what works in education. As such, we review studies on education
> interventions that may be included in our reports.
>
> The purpose of the attached letter is to notify you that we are in the
> process of conducting a review of High School Math interventions and
> may be including University of Chicago School Math Project in our
> review. The High School Math review focuses on mathematics
> interventions for high school students in grades 9-12 designed to
> impact student achievement, including curriculum-based interventions,
> instructional techniques, and products designed to deliver content and
> monitor student progress. In this letter, we ask you to review a list
> of studies, review a brief intervention summary, and sign an embargo
> agreement.
>
> We originally emailed for your input on May 12^th , 2010. The original
> deadline for a response has passed. We are trying to maintain a brief

> pace, and ask that you respond as soon as possible.
>
> Sincerely,
>
> Neil Seftor, Ph.D.
>
> Principal Investigator, WWC High School Math review
>
> <<UCSMP Developer Letter.pdf>>
>

From: WhatWorks
Sent: Monday, October 25, 2010 1:49 PM
To: 'z-usiskin@uchicago.edu'
Subject: RE: WWC High School Math Review of University of Chicago School Math Project (WWCPC 2237)
Dr. Usiskin,

We appreciate the time you took to respond to our request for information regarding UCSMP. Below, we have attempted to address the five major weaknesses of the WWC you describe in your response. It seems that many of them are based on misunderstandings regarding the WWC processes, so I have tried to clarify those below. Please feel free to follow-up with any specific concerns regarding WWC reports.

Sincerely,
Neil Seftor, Ph.D.
Principal Investigator, WWC High School Math Review

1. There is little if any conception of what constitutes a good sample for a typical curriculum study. Specifically, the WWC reviews do not seem to understand that the appropriate bite size for a curriculum study is not the student, but the classroom, for it is typically the classroom that is given a treatment. As a result, the reviewers look upon studies with one classroom (n = 1) and somehow think that this gives a reliable picture of how materials will fare in others. The naiveté here is that anyone who has implemented curriculum in multiple classrooms and in multiple schools knows that the same curriculum can fare well in one place and not so well in another. Teachers and school climate influence the implemented and learned curriculum in a wide variety of ways.

The result of this first weakness is that the WWC reviews include studies that are too small to be generalizable in any way and exclude large-scale studies that by their very size do not present the exact same conditions in the implementation of the materials but more accurately reflect the variety of ways in which materials are used and are more likely to have generalizable results.

> **The WWC evaluates studies of <u>all</u> sizes (larger than n=1) and provides <u>guidance on generalizability</u> through the extent of evidence categorization.**
>
> The WWC is focused on student-level outcomes, as are most studies of academic achievement. However, we agree that one classroom is not enough to provide evidence of effectiveness. Any study with only one classroom in a treatment or comparison condition would not meet WWC standards for review. See the discussion of confounding factors in the WWC Procedures and Standards Handbook.
>
> While it would indeed be preferable to have several large studies covering multiple settings, that is often not the case in educational research. The WWC is limited to reporting on existing studies that meet our standards, including small studies (with more than one classroom) as well as large ones; we do not exclude studies solely due to size. Furthermore, based on the number of studies and sizes of samples, we provide some information on how generalizable the results are using the extent of evidence categorization. It is intended as a companion to the effectiveness rating and should lead educators to interpret effectiveness ratings with caution when the extent of evidence is small. The WWC is

currently developing new dimensions for the extent of evidence classification that will make it more informative to readers.

2. There is no consideration of the treatment in the control group. There does not seem even to be any realization of what would seem to be obvious, that in a comparison study the comparative performance of two treatments depends on the quality of both treatments, not just the quality of the new treatment being examined.

Thus, for instance, if one is testing new materials in a school where the performance is historically low, there is more room for the new materials to perform better than the old than if the testing takes place in a school where the performance is historically high. And, in fact, since there are a great number of curricula in use, if there were some order to the quality of curricula, it is possible for a study to decide that curriculum X was not as effective as curriculum Y and not realize that curriculum X might be better than every other available curriculum.

Thus, to pass judgment on a curriculum on the basis of its performance against only one other curriculum can at best tell only a small part of the story of the relative effectiveness of the curriculum. If curriculum X is less effective than curriculum Y, X could be anywhere from the worst to the second best curriculum available. If it is more effective, it could be anywhere from the best to the second worst available.

The WWC evaluates <u>all</u> existing comparisons of an intervention and <u>documents the counterfactual</u> for each analysis.

We agree that it would be preferable to have an intervention evaluated against a number of other curricula. However, the mission of the WWC is to evaluate existing research and determine whether it meets our standards; that is, we do not have any control over the comparisons that have been made in the research literature. To the extent possible, we document the comparison condition in the description of each study, so that readers are aware of the environment in which the intervention was evaluated and the counterfactual against which it was compared.

3. The reviewers do not seem to be able to distinguish studies that are formative in nature - i.e., designed to inform the curriculum developers of how their materials are faring so that the materials can be improved before publication - and those that are summative in nature - i.e., with published materials. That is, the conditions under which studies are undertaken are usually not given any consideration.

The result of this third weakness is that the WWC reviews often consider studies of test and pilot editions of materials and ignore the existence of later versions of the same materials that may have taken into account the results of the very studies that are being reviewed.

The WWC evaluates <u>all</u> existing <u>effectiveness</u> studies of an intervention, regardless of version.

The WWC considers all studies that provide examine the effectiveness of a curriculum against an alternative. All such studies are eligible for review, regardless of version, and no studies are excluded based on later versions of materials. The versions of the interventions examined are documented in the report when they are described in the research. If there is reason to believe an intervention has changed substantially, we will investigate the

possibility further, and create separate reports if necessary. We do, however, exclude studies that solely compare different versions of a single intervention.

4. With respect to instruments, the reviewers seem to assume that all curricula are trying to cover the same content and that the differences are mainly instructional. With projects like ours, this is an invalid assumption - indeed, we would not exist if all we are trying to do is teach the old curriculum better. We are attempting to modernize the curriculum, to bring it up to date, to include new ideas and exclude old ones. The experiences our students have are quite different from the experiences of students in the comparable existing courses. When students have not had an opportunity to learn content (whether old or new), one has to be very careful about giving them a test over that content, because the student may simply give up and not even attempt the items that the student feels have not been covered. Furthermore, some of the content is so new that there are no standardized tests that cover it. For these reasons, we have to very carefully analyze test results with respect to opportunity to learn and we have to devise some of our own tests.

The result of this fourth weakness is that the WWC reviews of new materials often do not take into account how students do on the content that is new in these materials, that is, on the content that makes them the most different.

The WWC evaluates <u>all</u> valid and reliable outcome measures, including those which include <u>treatment-specific content</u>.

The WWC agrees that part of curriculum development may involve specific content changes.

5. An additional issue to consider when studying high school treatments is that beyond algebra I, there have not been current standardized tests that are content specific. That is, there have not been standardized tests focusing on geometry or algebra II or courses beyond algebra II. (We understand that Achieve is currently working to develop such a test at the algebra II level.) Thus, projects often have no opportunity to use a standardized instrument at these levels unless it is a general test, such as the SAT or the ITED-Q. At UCSMP, although we have developed our own tests, we have the classroom teachers indicate whether they taught or reviewed the content needed for students to answer the items and analyze the results based on teachers responses, both the UCSMP teachers and comparison teachers. We believe this is a transparent way to reduce bias. Yet, WWC completely discounts such studies. We believe this is unfortunate, particularly because we have data to indicate that there have often been many items on a standardized test, even at algebra I, for which teachers in both UCSMP and comparison treatments did not teach the necessary content.

The WWC evaluates <u>all</u> valid and reliable outcome measures and does not discount findings due to <u>implementation</u>.

As noted above, the WWC will include valid and reliable non-standardized outcomes, along with standardized measures, as researchers and policymakers may wish to glean information from a variety of sources. However, whether certain aspects of a curriculum were actually taught during the study period is more of a question of implementation and not the focus of WWC reviews. We will document implementation issues in our discussion of each study, but we will not exclude any study based on the specific content taught during the course of the study. If pieces of outcome measures are used based on the content taught, the study must demonstrate that those pieces must meet the validity and reliability requirements to be included in the intervention report.

From: Zalman Usiskin [z-usiskin@uchicago.edu]
Sent: Tuesday, November 09, 2010 3:45 PM
To: WhatWorks
Subject: Re: WWC High School Math Review of University of Chicago School Math Project (WWCPC 2237)

Follow Up Flag: Follow up
Flag Status: Red

Dear Dr. Seftor:

Thank you also for the time you took to respond to my e-mail.

I believe that you have misinterpreted the intent of my e-mail. I was not questioning whether you were following particular prescribed WWC processes or not; I was questioning the validity of the processes themselves to determine whether one curriculum is better than another. That you responded by reiterating WWC policies only serves to confirm that I am not misreading those policies and that my concerns are valid ones.

Basically, the WWC has tackled a job for which the data do not exist for any materials in the entire world. In the U.S., one possible way of collecting data on student performance vs. textbook usage would be to record the ISBN of the main text used by every student who takes the NAEP tests, and then to publish the distribution of scores by text used. But even these data would not be definitive, because certain texts may tend to be used in schools with higher-performing students and others with lower-performing students.

The WWC guidelines have been adopted as if a textbook is a pill and one can compare pill A to pill B with a randomized study. But textbooks are not pills. Once a pill is in the system, the pill-taker can do little to influence its effect on his or her system. However, once a textbook is adopted (i.e., in the classroom), both teacher and student have complete freedom to use, abuse, disuse, or misuse the book. In that kind of environment, studies have to be very sensitive to such things as the fidelity of the implementation of the text, the previous experience the teacher has with the material (e.g., don't necessarily test a book the first time the teacher goes through it), and the climate of the school and the class towards homework and achievement.

We believe in our UCSMP studies that we have done as good a job as anyone in providing a large-scale study in which we match classes at the beginning of the school year and examine overall scores on tests at the end of the year and also selected scores on collections of items that teachers indicate they covered during the year. It is very frustrating when studies that we view as well-controlled (whether involving UCSMP texts or not) are not considered as meeting WWC guidelines while studies we view as full of difficulties are used by WWC to make pronouncements that we feel are not warranted by the data.

As the book-length report, /On Evaluating Curricular Effectiveness/ (National Research Council, 2004), indicates, studies about curricular

effectiveness must “present evidence that provides reliable and valid indicators of the extent, quality, and type of the implementation of the materials. At a minimum, there should be documentation of the extent of coverage of curricular material” (p. 194). It appears that UCSMP studies of courses beyond geometry are often discounted solely because we use a project constructed test, despite reporting a KR20 value for the test, despite reporting information from the teacher about the appropriateness of various test items, and despite analyzing data based on teachers’ responses about the test items. So, it does appear that some studies are discounted based on implementation issues (e.g., implementation of a non-standardized posttest).

Finally, you may wish to take a look at the article “Myths about Curriculum Implementation”, by Denisse Thompson and Sharon Senk, in the NCTM 2010 Yearbook, /Mathematics Curriculum: Isses, Trends, and Future Directions/, pp. 249-263.

Zalman Usiskin

WhatWorks wrote:
> Dr. Usiskin,
>
> We appreciate the time you took to respond to our request for information regarding UCSMP. Below, we have attempted to address the five major weaknesses of the WWC you describe in your response. It seems that many of them are based on misunderstandings regarding the WWC processes, so I have tried to clarify those below. Please feel free to follow-up with any specific concerns regarding WWC reports.
>
> Sincerely,
>
> Neil Seftor, Ph.D.
>
> Principal Investigator, WWC High School Math Review
>
> /1. There is little if any conception of what constitutes a good sample for a typical curriculum study. Specifically, the WWC reviews do not seem to understand that the appropriate bite size for a curriculum study is not the student, but the classroom, for it is typically the classroom that is given a treatment. As a result, the reviewers look upon studies with one classroom (n = 1) and somehow think that this gives a reliable picture of how materials will fare in others. The naiveté here is that anyone who has implemented curriculum in multiple classrooms and in multiple schools knows that the same curriculum can fare well in one place and not so well in another. Teachers and school climate influence the implemented and learned curriculum in a wide variety of ways./
>
> / /
>
> /The result of this first weakness is that the WWC reviews include studies that are too small to be generalizable in any way and exclude

> large-scale studies that by their very size do not present the exact
> same conditions in the implementation of the materials but more
> accurately reflect the variety of ways in which materials are used and
> are more likely to have generalizable results./
>
> *The WWC evaluates studies of _all_ sizes (larger than n=1) and
> provides _guidance on generalizability_ through the extent of evidence
> categorization. *
>
> * *
>
> The WWC is focused on student-level outcomes, as are most studies of
> academic achievement. However, we agree that one classroom is not
> enough to provide evidence of effectiveness. Any study with only one
> classroom in a treatment or comparison condition would not meet WWC
> standards for review. See the discussion of confounding factors
> <http://ies.ed.gov/ncee/wwc/references/idocviewer/Doc.aspx?docId=19&to
> cId=4#confounding> in the WWC Procedures and Standards Handbook.
>
> While it would indeed be preferable to have several large studies
> covering multiple settings, that is often not the case in educational
> research. The WWC is limited to reporting on existing studies that
> meet our standards, including small studies (with more than one
> classroom) as well as large ones; we do not exclude studies solely due
> to size. Furthermore, based on the number of studies and sizes of
> samples, we provide some information on how generalizable the results
> are using the extent of evidence categorization. It is intended as a
> companion to the effectiveness rating and should lead educators to
> interpret effectiveness ratings with caution when the extent of
> evidence is small. The WWC is currently developing new dimensions for
> the extent of evidence classification that will make it more
> informative to readers.
>
> /2. There is no consideration of the treatment in the “control” group.
> There does not seem even to be any realization of what would seem to
> be obvious, that in a comparison study the comparative performance of
> two treatments depends on the quality of both treatments, not just the
> quality of the “new” treatment being examined./
>
> / /
>
> /Thus, for instance, if one is testing new materials in a school where
> the performance is historically low, there is more room for the new
> materials to perform better than the old than if the testing takes
> place in a school where the performance is historically high. And, in
> fact, since there are a great number of curricula in use, if there
> were some order to the quality of curricula, it is possible for a
> study to decide that curriculum X was not as effective as curriculum Y
> and not realize that curriculum X might be better than every other
> available curriculum./
>
> / /
>
> /Thus, to pass judgment on a curriculum on the basis of its

> performance against only one other curriculum can at best tell only a
> small part of the story of the relative effectiveness of the
> curriculum. If curriculum X is less effective than curriculum Y, X
> could be anywhere from the worst to the second best curriculum
> available. If it is more effective, it could be anywhere from the best
> to the second worst available./
>
> *The WWC evaluates _all_ existing comparisons of an intervention and
> _documents the counterfactual_ for each analysis.*
>
> We agree that it would be preferable to have an intervention evaluated
> against a number of other curricula. However, the mission of the WWC
> is to evaluate existing research and determine whether it meets our
> standards; that is, we do not have any control over the comparisons
> that have been made in the research literature. To the extent
> possible, we document the comparison condition in the description of
> each study, so that readers are aware of the environment in which the
> intervention was evaluated and the counterfactual against which it was
> compared.
>
> /3. The reviewers do not seem to be able to distinguish studies that
> are formative in nature - i.e., designed to inform the curriculum
> developers of how their materials are faring so that the materials can
> be improved before publication - and those that are summative in
> nature - i.e., with published materials. That is, the conditions under
> which studies are undertaken are usually not given any consideration./
>
> / /
>
> /The result of this third weakness is that the WWC reviews often
> consider studies of test and pilot editions of materials and ignore
> the existence of later versions of the same materials that may have
> taken into account the results of the very studies that are being
> reviewed./
>
> *The WWC evaluates _all_ existing _effectiveness_ studies of an
> intervention, regardless of version.*
>
> The WWC considers all studies that provide examine the effectiveness
> of a curriculum against an alternative. All such studies are eligible
> for review, regardless of version, and no studies are excluded based
> on later versions of materials. The versions of the interventions
> examined are documented in the report when they are described in the
> research. If there is reason to believe an intervention has changed
> substantially, we will investigate the possibility further, and create
> separate reports if necessary. We do, however, exclude studies that
> solely compare different versions of a single intervention.
>
> /4. With respect to instruments, the reviewers seem to assume that all
> curricula are trying to cover the same content and that the
> differences are mainly instructional. With projects like ours, this is
> an invalid assumption - indeed, we would not exist if all we are
> trying to do is teach the old curriculum better. We are attempting to
> modernize the curriculum, to bring it up to date, to include new ideas

> and exclude old ones. The experiences our students have are quite
> different from the experiences of students in the comparable existing
> courses. When students have not had an opportunity to learn content
> (whether old or new), one has to be very careful about giving them a
> test over that content, because the student may simply give up and not
> even attempt the items that the student feels have not been covered.
> Furthermore, some of the content is so new that there are no
> standardized tests that cover it. For these reasons, we have to very
> carefully analyze test results with respect to opportunity to learn
> and we have to devise some of our own tests./
>
> / /
>
> /The result of this fourth weakness is that the WWC reviews of new
> materials often do not take into account how students do on the
> content that is new in these materials, that is, on the content that
> makes them the most different./
>
> *The WWC evaluates _all_ valid and reliable outcome measures,
> including those which include _treatment-specific content_.*
>
> The WWC agrees that part of curriculum development may involve
> specific content changes.
>
> /5. An additional issue to consider when studying high school
> treatments is that beyond algebra I, there have not been current
> standardized tests that are content specific. That is, there have not
> been standardized tests focusing on geometry or algebra II or courses
> beyond algebra II. (We understand that Achieve is currently working to
> develop such a test at the algebra II level.) Thus, projects often
> have no opportunity to use a standardized instrument at these levels
> unless it is a general test, such as the SAT or the ITED-Q. At UCSMP,
> although we have developed our own tests, we have the classroom
> teachers indicate whether they taught or reviewed the content needed
> for students to answer the items and analyze the results based on
> teachers' responses, both the UCSMP teachers and comparison teachers.
> We believe this is a transparent way to reduce bias. Yet, WWC
> completely discounts such studies. We believe this is unfortunate,
> particularly because we have data to indicate that there have often
> been many items on a standardized test, even at algebra I, for which
> teachers in both UCSMP and comparison treatments did not teach the
> necessary content./
>
> *The WWC evaluates _all_ valid and reliable outcome measures and does
> not discount findings due to _implementation_.*
>
> As noted above, the WWC will include valid and reliable
> non-standardized outcomes, along with standardized measures, as
> researchers and policymakers may wish to glean information from a
> variety of sources. However, whether certain aspects of a curriculum
> were actually taught during the study period is more of a question of
> implementation and not the focus of WWC reviews. We will document
> implementation issues in our discussion of each study, but we will not
> exclude any study based on the specific content taught during the

> course of the study. If pieces of outcome measures are used based on
> the content taught, the study must demonstrate that those pieces must
> meet the validity and reliability requirements to be included in the
> intervention report.
>

From: WhatWorks
Sent: Wednesday, February 16, 2011 11:27 AM
To: 'Zalman Usiskin'
Subject: RE: WWC High School Math Review of University of Chicago School Math Project (WWCPC 2237)

Dr. Usiskin,

Thank you for your thoughtful email considering WWC reviews. Below, I have attempted to address the issues you describe in your response. Please feel free to follow-up with any concerns regarding specific WWC reports.

I believe that you have misinterpreted the intent of my e-mail. I was not questioning whether you were following particular prescribed WWC processes or not; I was questioning the validity of the processes themselves to determine whether one curriculum is better than another.

> WWC reviews do not directly compare curricula; however, decision-makers may use existing research to compare curricula. Some of that research is reliable and some is not, so the goal of the WWC is to identify and summarize the findings of the most reliable research. Decision-makers can then make the same comparisons they would make anyway, including cross-curricular comparisons, but base them on the reliable research.

That you responded by reiterating WWC policies only serves to confirm that I am not misreading those policies and that my concerns are valid ones.

> The policies were described to illustrate that they actually agree with what you were suggesting. Therefore, if you notice discrepancies, we should investigate and correct them as appropriate.

It is very frustrating when studies that we view as well-controlled (whether involving UCSMP texts or not) are not considered as meeting WWC guidelines while studies we view as full of difficulties are used by WWC to make pronouncements that we feel are not warranted by the data.

It appears that UCSMP studies of courses beyond geometry are often discounted solely because we use a project constructed test, despite reporting a KR20 value for the test, despite reporting information from the teacher about the appropriateness of various test items, and despite an*alyzing data based on teachers' responses about the test items.*

> If you have examples of either of these issues, please let us know which studies and WWC reports are involved so that we may investigate.

Sincerely,

Neil Seftor, Ph.D.
Principal Investigator, WWC High School Math Review

-----Original Message-----
From: Zalman Usiskin [mailto:z-usiskin@uchicago.edu]
Sent: Tuesday, November 09, 2010 3:45 PM
To: WhatWorks
Subject: Re: WWC High School Math Review of University of Chicago School Math Project (WWCPC 2237)

Dear Dr. Seftor:

Thank you also for the time you took to respond to my e-mail.

I believe that you have misinterpreted the intent of my e-mail. I was not questioning whether you were following particular prescribed WWC processes or not; I was questioning the validity of the processes themselves to determine whether one curriculum is better than another.
That you responded by reiterating WWC policies only serves to confirm that I am not misreading those policies and that my concerns are valid ones.

Basically, the WWC has tackled a job for which the data do not exist for any materials in the entire world. In the U.S., one possible way of collecting data on student performance vs. textbook usage would be to record the ISBN of the main text used by every student who takes the NAEP tests, and then to publish the distribution of scores by text used.
But even these data would not be definitive, because certain texts may tend to be used in schools with higher-performing students and others with lower-performing students.

The WWC guidelines have been adopted as if a textbook is a pill and one can compare pill A to pill B with a randomized study. But textbooks are not pills. Once a pill is in the system, the pill-taker can do little to influence its effect on his or her system. However, once a textbook is adopted (i.e., in the classroom), both teacher and student have complete freedom to use, abuse, disuse, or misuse the book. In that kind of environment, studies have to be very sensitive to such things as the fidelity of the implementation of the text, the previous experience the teacher has with the material (e.g., don't necessarily test a book the first time the teacher goes through it), and the climate of the school and the class towards homework and achievement.

We believe in our UCSMP studies that we have done as good a job as anyone in providing a large-scale study in which we match classes at the beginning of the school year and examine overall scores on tests at the end of the year and also selected scores on collections of items that teachers indicate they covered during the year. It is very frustrating when studies that we view as well-controlled (whether involving UCSMP texts or not) are not considered as meeting WWC guidelines while studies we view as full of difficulties are used by WWC to make pronouncements that we feel are not warranted by the data.

As the book-length report, /On Evaluating Curricular Effectiveness/ (National Research Council, 2004), indicates, studies about curricular effectiveness must "present evidence that provides reliable and valid indicators of the extent, quality, and type of the implementation of the materials. At a minimum, there should be documentation of the extent of coverage of curricular material" (p. 194). It appears that UCSMP studies of courses beyond geometry are often discounted solely because we use a project

constructed test, despite reporting a KR20 value for the test, despite reporting information from the teacher about the appropriateness of various test items, and despite analyzing data based on teachers' responses about the test items. So, it does appear that some studies are discounted based on implementation issues (e.g., implementation of a non-standardized posttest).

Finally, you may wish to take a look at the article "Myths about Curriculum Implementation", by Denisse Thompson and Sharon Senk, in the NCTM 2010 Yearbook, /Mathematics Curriculum: Isses, Trends, and Future Directions/, pp. 249-263.

Zalman Usiskin

WhatWorks wrote:
>
> Dr. Usiskin,
>
> We appreciate the time you took to respond to our request for
> information regarding UCSMP. Below, we have attempted to address the
> five major weaknesses of the WWC you describe in your response. It
> seems that many of them are based on misunderstandings regarding the
> WWC processes, so I have tried to clarify those below. Please feel
> free to follow-up with any specific concerns regarding WWC reports.
>
> Sincerely,
>
> Neil Seftor, Ph.D.
>
> Principal Investigator, WWC High School Math Review
>
> /1. There is little if any conception of what constitutes a good
> sample for a typical curriculum study. Specifically, the WWC reviews
> do not seem to understand that the appropriate bite size for a
> curriculum study is not the student, but the classroom, for it is
> typically the classroom that is given a treatment. As a result, the
> reviewers look upon studies with one classroom (n = 1) and somehow
> think that this gives a reliable picture of how materials will fare in
> others. The naiveté here is that anyone who has implemented curriculum
> in multiple classrooms and in multiple schools knows that the same
> curriculum can fare well in one place and not so well in another.
> Teachers and school climate influence the implemented and learned
> curriculum in a wide variety of ways./
>
> / /
>
> /The result of this first weakness is that the WWC reviews include
> studies that are too small to be generalizable in any way and exclude
> large-scale studies that by their very size do not present the exact
> same conditions in the implementation of the materials but more

> accurately reflect the variety of ways in which materials are used and
> are more likely to have generalizable results./
>
> *The WWC evaluates studies of _all_ sizes (larger than n=1) and
> provides _guidance on generalizability_ through the extent of evidence
> categorization. *
>
> * *
>
> The WWC is focused on student-level outcomes, as are most studies of
> academic achievement. However, we agree that one classroom is not
> enough to provide evidence of effectiveness. Any study with only one
> classroom in a treatment or comparison condition would not meet WWC
> standards for review. See the discussion of confounding factors
> <http://ies.ed.gov/ncee/wwc/references/idocviewer/Doc.aspx?docId=19&to
> cId=4#confounding> in the WWC Procedures and Standards Handbook.
>
> While it would indeed be preferable to have several large studies
> covering multiple settings, that is often not the case in educational
> research. The WWC is limited to reporting on existing studies that
> meet our standards, including small studies (with more than one
> classroom) as well as large ones; we do not exclude studies solely due
> to size. Furthermore, based on the number of studies and sizes of
> samples, we provide some information on how generalizable the results
> are using the extent of evidence categorization. It is intended as a
> companion to the effectiveness rating and should lead educators to
> interpret effectiveness ratings with caution when the extent of
> evidence is small. The WWC is currently developing new dimensions for
> the extent of evidence classification that will make it more
> informative to readers.
>
> /2. There is no consideration of the treatment in the “control” group.
> There does not seem even to be any realization of what would seem to
> be obvious, that in a comparison study the comparative performance of
> two treatments depends on the quality of both treatments, not just the
> quality of the “new” treatment being examined./
>
> / /
>
> /Thus, for instance, if one is testing new materials in a school where
> the performance is historically low, there is more room for the new
> materials to perform better than the old than if the testing takes
> place in a school where the performance is historically high. And, in
> fact, since there are a great number of curricula in use, if there
> were some order to the quality of curricula, it is possible for a
> study to decide that curriculum X was not as effective as curriculum Y
> and not realize that curriculum X might be better than every other
> available curriculum./

>
> / /
>
> /Thus, to pass judgment on a curriculum on the basis of its
> performance against only one other curriculum can at best tell only a
> small part of the story of the relative effectiveness of the
> curriculum. If curriculum X is less effective than curriculum Y, X
> could be anywhere from the worst to the second best curriculum
> available. If it is more effective, it could be anywhere from the best
> to the second worst available./
>
> *The WWC evaluates _all_ existing comparisons of an intervention and
> _documents the counterfactual_ for each analysis.*
>
> We agree that it would be preferable to have an intervention evaluated
> against a number of other curricula. However, the mission of the WWC
> is to evaluate existing research and determine whether it meets our
> standards; that is, we do not have any control over the comparisons
> that have been made in the research literature. To the extent
> possible, we document the comparison condition in the description of
> each study, so that readers are aware of the environment in which the
> intervention was evaluated and the counterfactual against which it was
> compared.
>
> /3. The reviewers do not seem to be able to distinguish studies that
> are formative in nature - i.e., designed to inform the curriculum
> developers of how their materials are faring so that the materials can
> be improved before publication - and those that are summative in
> nature - i.e., with published materials. That is, the conditions under
> which studies are undertaken are usually not given any consideration./
>
> / /
>
> /The result of this third weakness is that the WWC reviews often
> consider studies of test and pilot editions of materials and ignore
> the existence of later versions of the same materials that may have
> taken into account the results of the very studies that are being
> reviewed./
>
> *The WWC evaluates _all_ existing _effectiveness_ studies of an
> intervention, regardless of version.*
>
> The WWC considers all studies that provide examine the effectiveness
> of a curriculum against an alternative. All such studies are eligible
> for review, regardless of version, and no studies are excluded based
> on later versions of materials. The versions of the interventions
> examined are documented in the report when they are described in the
> research. If there is reason to believe an intervention has changed

> substantially, we will investigate the possibility further, and create
> separate reports if necessary. We do, however, exclude studies that
> solely compare different versions of a single intervention.
>
> /4. With respect to instruments, the reviewers seem to assume that all
> curricula are trying to cover the same content and that the
> differences are mainly instructional. With projects like ours, this is
> an invalid assumption - indeed, we would not exist if all we are
> trying to do is teach the old curriculum better. We are attempting to
> modernize the curriculum, to bring it up to date, to include new ideas
> and exclude old ones. The experiences our students have are quite
> different from the experiences of students in the comparable existing
> courses. When students have not had an opportunity to learn content
> (whether old or new), one has to be very careful about giving them a
> test over that content, because the student may simply give up and not
> even attempt the items that the student feels have not been covered.
> Furthermore, some of the content is so new that there are no
> standardized tests that cover it. For these reasons, we have to very
> carefully analyze test results with respect to opportunity to learn
> and we have to devise some of our own tests./
>
> / /
>
> /The result of this fourth weakness is that the WWC reviews of new
> materials often do not take into account how students do on the
> content that is new in these materials, that is, on the content that
> makes them the most different./
>
> *The WWC evaluates _all_ valid and reliable outcome measures,
> including those which include _treatment-specific content_.*
>
> The WWC agrees that part of curriculum development may involve
> specific content changes.
>
> /5. An additional issue to consider when studying high school
> treatments is that beyond algebra I, there have not been current
> standardized tests that are content specific. That is, there have not
> been standardized tests focusing on geometry or algebra II or courses
> beyond algebra II. (We understand that Achieve is currently working to
> develop such a test at the algebra II level.) Thus, projects often
> have no opportunity to use a standardized instrument at these levels
> unless it is a general test, such as the SAT or the ITED-Q. At UCSMP,
> although we have developed our own tests, we have the classroom
> teachers indicate whether they taught or reviewed the content needed
> for students to answer the items and analyze the results based on
> teachers' responses, both the UCSMP teachers and comparison teachers.
> We believe this is a transparent way to reduce bias. Yet, WWC
> completely discounts such studies. We believe this is unfortunate,

> particularly because we have data to indicate that there have often
> been many items on a standardized test, even at algebra I, for which
> teachers in both UCSMP and comparison treatments did not teach the
> necessary content./
>
> *The WWC evaluates _all_ valid and reliable outcome measures and does
> not discount findings due to _implementation_.*
>
> As noted above, the WWC will include valid and reliable
> non-standardized outcomes, along with standardized measures, as
> researchers and policymakers may wish to glean information from a
> variety of sources. However, whether certain aspects of a curriculum
> were actually taught during the study period is more of a question of
> implementation and not the focus of WWC reviews. We will document
> implementation issues in our discussion of each study, but we will not
> exclude any study based on the specific content taught during the
> course of the study. If pieces of outcome measures are used based on
> the content taught, the study must demonstrate that those pieces must
> meet the validity and reliability requirements to be included in the
> intervention report.
>

From: What Works
Sent: Thursday, September 29, 2011 4:20 PM
To: 'Zalman Usiskin'
Subject: What Works Clearinghouse (WWCPC 2011002)
Attachments: QRT 2011002.pdf

Dear Dr. Usiskin,

Attached is a response to the questions you raised in your July 21 message to the What Works Clearinghouse (WWC).

Thank you,

What Works Clearinghouse

From: WhatWorks
Sent: Wednesday, August 10, 2011 4:20 PM
To: 'Zalman Usiskin'
Subject: RE: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Dr. Usiskin,

The WWC Quality Review Team is reviewing your email and will prepare a written response. The Quality Review Team responds to concerns raised by study authors, curriculum developers or other relevant parties about WWC reviews published on our website. These quality reviews are undertaken when concerned parties present evidence that a WWC review may be inaccurate. When a quality review is conducted, a researcher who was not involved in the initial review undertakes an independent assessment of the study in question. The researcher also investigates the procedures used and decisions made during the original review of the study. If a quality review concludes that the original review contained errors, a revision will be published. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the What Works Clearinghouse.

Thank you,

What Works Clearinghouse

From: WhatWorks
Sent: Friday, July 22, 2011 2:18 PM
To: 'Zalman Usiskin'
Subject: RE: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Dr. Usiskin,

Thank you for contacting the What Works Clearinghouse (WWC). We have received your email below. WWC staff are reviewing your request and will prepare a response.

What Works Clearinghouse

The What Works Clearinghouse was established by the U.S. Department of Education's Institute of Education Sciences to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education. For more information, please visit http://ies.ed.gov/ncee/wwc/.

-----Original Message-----
From: Zalman Usiskin [mailto:z-usiskin@uchicago.edu]
Sent: Thursday, July 21, 2011 2:18 PM
To: What Works
Cc: denisse@usf.edu
Subject: Re: WWC High School Math Review of University of Chicago School Mathematics Project 6-12 Curriculum

Dear Ms. Constantine:

Thank you for this notification.

We appreciate the clarity and completeness of your overall report and we appreciate your acceptance of two studies as meeting WWC criteria.

We think, however, that you have mistakenly excluded two others. The reports of the evaluations of UCSMP Advanced Algebra (Thompson, Senk, Witonsky, Usiskin, & Kaeley, 2001) and UCSMP Geometry (Thompson, Witonsky, Senk, Usiskin, & Kaeley, 2003) each contain a detailed description of the strict procedures used to determine whether supposedly matched pairs of treatment and comparison classes were actually equivalent at the beginning of the year and to eliminate from the analysis those pairs of classes that were not. You might wish to look again at the reports of these studies.

Zalman Usiskin
Professor Emeritus of Education
Director, UCSMP
The University of Chicago
1225 E. 60th St.
Chicago, IL 60637

What Works wrote:

>
> Dear Dr. Usiskin and Dr. Thompson,
>
>
>
> The attached letter is to notify you that the What Works Clearinghouse
> (WWC) has completed the review of the research on /University of
> Chicago School Mathematics Project (UCSMP) 6-12 Curriculum/ and
> determined that this intervention is eligible for an intervention
> report according to the High School Math review protocol. We have
> also attached a courtesy copy of the report which will be posted on
> the WWC website on July 19, 2011. As a reminder, this report is
> covered by the embargo agreement sent to you on May 12, 2010,
> requiring you not to copy, distribute, or discuss the report with
> members of the public outside your organization, prior to release of
> the report by the Institute of Education Sciences.
>
>
>
> Sincerely,
>
>
>
> Jill Constantine
>
> Director, What Works Clearinghouse
>
>
>

September 29, 2011

Professor Zalman Usiskin
Director, UCSMP
The University of Chicago
1225 E. 60th Street
Chicago, IL 60637
z-usiskin@uchicago.edu

Reference: QR2011002

Dear Professor Usiskin:

Thank you for your inquiry concerning the WWC intervention report on the University of Chicago School Mathematics Project (UCSMP). In response to your email, we conducted an independent quality review. The WWC quality review team responds to concerns raised by study authors, curriculum developers, or other relevant parties about WWC reviews published on our website. When a quality review is conducted, a researcher who was not involved in the initial review undertakes an independent assessment of the studies in question. The researcher also investigates the procedures used and decisions made during the original review of the studies. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the WWC.

Your email suggested that the WWC reconsider two reports: *An evaluation of the second edition of UCSMP Advanced Algebra* (Thompson, Senk, Witonsky, Usiskin, & Kaeley, 2001) and *An evaluation of the second edition of UCSMP Geometry* (Thompson, Witonsky, Senk, Usiskin, & Kaeley, 2003). For each of these studies, the WWC reported, "The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent." You noted that each of the studies contains a detailed description of the procedures used to determine equivalence and to eliminate pairs that were not equivalent.

The quality review team found that the studies do examine equivalence based on a pretest. Furthermore, as your email noted, pairs that showed statistically significant pretest differences were removed from the analysis. However, WWC standards for equivalence are based on the magnitude of the pretest differences relative to the standard deviation of the pretest, regardless of the statistical significance of the pretest differences. Specifically, the difference in the mean pretest between the analytic treatment and intervention groups is divided by the standard deviation (in the pooled sample). The resulting value, known as an "effect size," must be less than 0.25. In addition, if the value is greater than 0.05, the estimated impacts must be adjusted for baseline differences (for example, by using a regression adjustment that controls for pretest results). More information on

WWC standards for establishing equivalence can be found in the *WWC Procedures and Standards Handbook* (available at http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf, see page 15).

The quality review team verified that the original WWC reviewers calculated the effect sizes for pretest equivalence. Specifically, the WWC was most interested in the pooled results using the non-UCSMP comparison group. Based on the pretest statistics reported in the studies (Table 9 in the advanced algebra study and Table 8 in the geometry study), the original WWC reviewers calculated effect sizes for Part I and Part II of the pretests that were above 0.05 for the pooled samples (pools of 8 matched pairs in each study; Part I and Part II were treated separately). Based on these calculations, the original WWC reviewers concluded that the analytic samples in the studies did not meet WWC standards for pretest equivalence (effect sizes were above 0.05 and the study authors did not control for the pretest differences in their analyses). The quality review team conducted independent calculations that verified the conclusions of the original WWC reviewers.

I hope that this letter has addressed your concerns. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov.

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

cc: (b)(6)