

From: smetzger@mtrace.org
Sent: Wednesday, April 14, 2010 11:09 AM
To: info@whatworks.ed.gov
Subject: IES Website: Contact Us: Topic Areas, Beginning Reading
Review, Reference
ID Number: 1852710946

info@whatworks.ed.gov, this email was automatically sent through the
Contact
link on the WWC website.

From: smetzger@mtrace.org

Message: While I received the WWC's evidence standards, I continue to be perplexed about it's ratings for Reading Recovery. I have read the four studies that met the "standards." Two of them were conducted by Pinnell, who is introduced the program in the U.S. One of those studies compares the program to itself, with the only variable being the training of classroom teachers! The other study simply reveals that students in Reading Recovery performed better on the Observation Survey (created by Reading Recovery promoters) and Woodcock Johnson than students who received no intervention AT ALL. I just do not understand how the WWC can conclude that the program is effective based on the studies that met its standards. What is the WWC's view on the importance of ruling out bias on the part of the researchers? Please reply using my email shown above.

From: WhatWorks

Sent: Thursday, April 15, 2010 2:17 PM

To: 'smetzger@mtrace.org'

Subject: What Works Clearinghouse (WWCPC 2089)

Hello,

Thank you for contacting the What Works Clearinghouse (WWC). 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 was flawed, a revision will be published. These quality reviews are one of tools used to ensure that the standards established by the Institute of Educational Sciences (IES) are upheld on every review conducted by the What Works Clearinghouse.

Thank you,

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: smetzger@mtrace.org [mailto:smetzger@mtrace.org]

Sent: Wednesday, April 14, 2010 11:09 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Topic Areas, Beginning Reading Review,

Reference ID Number: 1852710946

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: smetzger@mtrace.org

Message: While I received the WWC's evidence standards, I continue to be perplexed about it's ratings for Reading Recovery. I have read the four studies that met the "standards." Two of them were conducted by Pinnell, who is introduced the program in the U.S. One of those studies compares the program to itself, with the only variable being the training of classroom teachers! The other study simply reveals that students in Reading Recovery performed better on the Observation Survey (created by Reading Recovery promoters) and Woodcock Johnson than students who received no

intervention AT ALL. I just do not understand how the WWC can conclude that the program is effective based on the studies that met its standards. What is the WWC's view on the importance of ruling out bias on the part of the researchers? Please reply using my email shown above.

Reading Recovery Receives High Rating for Effectiveness – Again

July 16, 2010

Reading Recovery, an early literacy intervention for first graders and a teacher professional development program, has once again received high research ratings from a federally funded information agency. The National Center on Response to Intervention (NCRTI) (<http://www.rti4success.org>) yesterday listed Reading Recovery on its intervention site and reported large gains based on a 2005 study by researcher Robert Schwartz published in the *Journal of Educational Psychology*. NCRTI is funded by the U.S. Department of Education's Office of Special Education Programs to provide technical assistance to states and districts to help them implement proven models for response to intervention (RTI) and early intervening services.

The new NCRTI ratings complement Reading Recovery's high marks from another independent rating agency, the USDE's What Works Clearinghouse (WWC). In 2007 and 2008, WWC awarded high ratings across four areas of reading including alphabetics, fluency, comprehension, and general reading achievement. To date, Reading Recovery remains the only early reading program to receive "positive" or "potentially positive" ratings in all areas. Although both WWC and NCRTI evaluate research on effectiveness, each has established its own criteria.

Reading Recovery as Response to Intervention

"This rating from NCRTI is important for curriculum decisionmakers because response to intervention is expanding in schools across the United States," said Jady Johnson, executive director of the Reading Recovery Council of North America. "The 2004 reauthorization of the Individuals with Disabilities Act included language to establish RTI as a way to intervene early with students at risk of failure and reduce inappropriate identification of children for special education."

With more than 25 years of implementation in the United States, Reading Recovery is a proven intervention used by thousands of elementary schools. Approximately 75% of children who complete the Reading Recovery intervention meet grade-level standards and continue to learn in classrooms without supplemental support. Since its 1984 introduction in the United States, nearly two million students have had Reading Recovery lessons.

The goal of Reading Recovery is to dramatically reduce the number of first-grade students who have extreme difficulty learning to read and write, and to reduce the cost of these learners to education systems. More information about Reading Recovery and RTI is available online at <http://fdf.readingrecovery.org/index.php/response-to-intervention>.

From: (b)(6)
Sent: Monday, May 24, 2010 9:55 AM
To: info@whatworks.ed.gov
Subject: IES Website: Contact Us: Evidence Standards, Reference ID
Number:
1415747212

info@whatworks.ed.gov, this email was automatically sent through the
Contact
link on the WWC website.

From: (b)(6)
Message: http://www.nrrf.org/rr_study_chapman.htm

From: WhatWorks

Sent: Monday, May 24, 2010 4:52 PM

To: (b)(6)

Subject: What Works Clearinghouse (WWCPC 2177)

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number: 1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: Stephanie Metzger (b)(6)

Sent: Saturday, June 05, 2010 10:16 AM

To: WhatWorks

Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10**, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>

Subject: What Works Clearinghouse (WWCPC 2177)

To: (b)(6)

Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: 

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From:

(b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: WhatWorks

Sent: Tuesday, June 08, 2010 11:10 AM

To: 'Stephanie Metzger'

Subject: RE: What Works Clearinghouse (WWCPC 2177)

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the WWC Procedures and Standards Handbook, inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [REDACTED]

Sent: Saturday, June 05, 2010 10:16 AM

To: WhatWorks

Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turnner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10**, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: [REDACTED]
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

This article was downloaded by:[University of Wollongong]
[University of Wollongong]

On: 18 June 2007

Access Details: [subscription number 773326199]

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954

Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Disability, Development and Education

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713425407>

Reading Recovery 20 Years Down the Track: Looking forward, looking back

To cite this Article: Reynolds, Mere and Wheldall, Kevin, 'Reading Recovery 20 Years Down the Track: Looking forward, looking back', International Journal of Disability, Development and Education, 54:2, 199 - 223

To link to this article: DOI: 10.1080/10349120701330503

URL: <http://dx.doi.org/10.1080/10349120701330503>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

© Taylor and Francis 2007

Reading Recovery 20 Years Down the Track: Looking forward, looking back

Meree Reynolds and Kevin Wheldall*

Macquarie University, Australia

Reading Recovery is an intensive literacy programme designed for young students who have been identified as being at-risk of reading failure after 1 year of schooling. The intervention was developed and trialled in New Zealand over 20 years ago and is now implemented in a number of education systems. The focus of this article is on recent research into the operationalisation of the programme with an overview of what it has done well and what it has not done so well. Reading Recovery has been very successful in bringing about change on the political and teacher training levels. In terms of efficacy in remediating literacy difficulties, however, the findings are more equivocal. What we have learned from Reading Recovery may assist in the implementation of new interventions based on more contemporary research.

Keywords: *Early intervention; Early literacy; Reading acquisition; Reading Recovery*

Introduction

Reading Recovery (RR), an early literacy preventative programme for at-risk students after 1 year of schooling, has been widely reported to be an effective intervention (Askew & Frasier, 1997; Johnston & Allington, 1991; Lyons, 1997; Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994; Reutzel, 1999; Snow, Burns, & Griffin, 1998). Its goal is to assist the lowest performing students in a school system after 1 year of schooling to improve to the average reading level of their peers within 12–20 weeks of intensive individual instruction (Clay, 1987), and, consequently, to reduce considerably the number of students requiring remedial and special education support in later years (Clay; Lyons).

Developed by Dame Marie Clay in New Zealand in the 1970s, RR has been implemented there as a system intervention for over 20 years. It has also been

*Corresponding author. Macquarie University Special Education Centre, Macquarie University, Sydney, NSW 2109, Australia. Email: kevin.wheldall@mq.edu.au

adopted by school systems in the United States (U.S.), Canada, Great Britain, and Australia for periods of between 10 and 20 years. It has been so successful in its breadth of implementation that, in the U.S. alone, more than 1.4 million students have been involved in the programme during 20 years of implementation (Reading Recovery Council of North America, 2004a).

As implementation of an educational intervention over a period of 20 years is a rare occurrence, it could be said that RR has achieved remarkable longevity. This raises the question of why, when educational innovations typically come and go relatively quickly, is RR still being implemented in many school systems? Is RR still being implemented because the targeted students are so successful? Is it still being implemented because it has proven cost benefits to an education system?

The programme has attracted considerable attention over the past two decades and there is a significant amount of data now available from evaluations. While many are very supportive of the programme and are enthusiastic about its merits (Askew et al., 2003; Lyons, 1997; Pinnell et al., 1994; Swartz & Klein, 1997), there have been a number of studies and articles that question whether the programme is as effective as it is claimed to be (Center, Wheldall, Freeman, Outhred, & McNaught, 1995; Shanahan & Barr, 1995). Others suggest that, as there are methodological weaknesses in much of the research, it is hard to draw conclusions about its effectiveness (Hiebert, 1994; Shanahan & Barr).

RR is designed for implementation within an education system. Clay suggests that there are four “dimensions” of change that need to be in place to facilitate its implementation within a system: “behavioural change on the part of teachers; child behaviour change achieved by teaching; organisational changes in schools achieved by teachers and administrators; and social/political changes in funding by controlling authorities” (Clay, 1987, p. 36).

This literature review will look back at the implementation of RR, focusing on data and recent research findings from studies in the past 15 years about the implementation and operationalisation of the RR programme in various school systems. This time period represents the time that has elapsed since Center, Wheldall, and Freeman (1992) completed a review of findings about RR. The current review will provide information about research studies and data pertaining to RR that have been published from 1992 to the present to determine whether there are common findings and trends. In addition, this article will focus on the four levels of implementation that Clay (1987) suggests must be in place in order to identify reasons for the longevity of the programme. The review will also look forward, drawing implications for future research and implementation of preventative programmes.

The Reading Recovery Programme

The RR programme provides intensive, one-to-one, daily tutoring for young children who are identified as being at-risk of having literacy difficulties after having received a full year of schooling. Students who are targeted for RR are the lowest performing students in the school as judged by the Observation Survey of Early

Literacy Achievement (Clay, 1993a), which has six components: a running record on text reading, letter identification, dictation, concepts about print, sight words, and writing vocabulary.

In each 30-min daily session a trained RR teacher carries out a number of set activities that are related to texts selected for the student's reading level. These activities are re-reading one or more previously introduced texts, identifying letters and words, writing a story, hearing and writing sounds in words, cutting the story up and then reassembling and reading it, introducing a new book, and reading the new text (Clay, 1993b).

Students are discontinued from the programme when they are able to read texts that an average reader in the child's class can read, can write several sentences (Clay, 1993b), and are "predicted to make progress without further individual instruction" (Lyons, 2003, p. 219). Thus, the length of the programme varies from student to student, usually taking between 12 and 16 weeks (Kerslake, 1999).

RR was designed as a secondary intervention programme that follows a primary prevention programme in which there is quality literacy teaching in the first year of schooling (Askew et al., 2003). The focus of the teaching programme is for targeted students to "accelerate" and make sufficient progress so that they catch up with the average student in the class.

Implementation within a System

RR has been promoted as a system intervention that is effective in reducing the demand for special education and remedial programmes in later years (Askew et al., 2003). In a school system, between 10% and 20% of Year 1 students may be at-risk and, therefore, may benefit from RR intervention (Shanahan & Barr, 1995). Clay summed up the potential of the programme in saying that, once the programme was established, students in the programme would no longer need intervention except "for the 1–2% of children who cannot be discontinued and special education groups" (Clay, 1998, p. 19). The programme's effectiveness as a system intervention, therefore, can be determined by searching for evidence that the demand for tertiary support services has lessened as RR has reached full implementation in a system.

The programme relies on knowledgeable, experienced, and well-trained teachers who interact with students using RR techniques with the aim that students acquire strategies for self-improvement and monitoring their own reading behaviour (Reynolds, 1993). The training programme and on-going professional development for teachers and teacher trainers are key components of the programme and are primarily based on Clay's publications and those of key RR personnel.

Schools have to make organisational changes to accommodate the programme. Depending upon the school system, this may mean providing additional funding, or teaching spaces and resources, rearranging staffing, and providing additional teacher time for assessment and monitoring. This is a condition of schools being involved in the programme (Clay, 1987).

RR is relatively costly compared with other interventions (Groff, n.d.; Hiebert, 1994; Shanahan & Barr, 1995). For example, Hiebert calculated that, at 1994 costs, the cost per successful student could be more than US\$8,000 (equivalent to approximately \$AU11,000). In 2003, the cost per RR student was calculated at AU\$9,088 (equivalent to US\$6,603) in the state of New South Wales (NSW) in Australia (NSW Department of Education and Training, 2003). While some individual schools may have funds to cover the cost of the teacher's salary and books, system funding is usually required for the salary of the tutors or teacher trainers, the establishment of training facilities, and, in some systems, salaries of the RR teachers.

RR requires a commitment from a system, a school district, or a government or government body for funding to be allocated. This commitment is the product of a belief that the programme is worth the cost: that it is so effective it will reduce the demand on costly special education and remedial services in later years.

RR administrators gather data on every student in the programme in a school system. This is used to monitor the effectiveness of the teaching programme for children by looking at the number and percentage of students who are discontinued. As RR has been established as a system intervention for over 10 years in several countries, there is now a significant amount of data available to measure the success of the teaching programme using RR's monitoring mechanism and other available system data.

Reading Recovery in New Zealand

RR was developed and trialled in New Zealand in the 1970s. The programme was designed to operate in New Zealand schools, and its procedures and teaching strategies are aligned with the literacy curriculum and the predominantly whole-language approach that is used in schools throughout the country (Tunmer & Chapman, 2003). Field trials, replications, and follow-up studies conducted by Clay showed that students who were successfully discontinued had demonstrated accelerated progress (Clay, 1993b). Subsequently, the programme was expanded in 1980/81, and in 1983 it was supported as a national programme with implementation in each administrative district. From the mid-1980s to the mid-1990s RR grew rapidly, and a system of collection and monitoring of national data was initiated in 1984.

National data for 1984–2003 show that approximately 60% of all students who entered RR were discontinued during the year they started in the programme (Anand & Bennie, 2004, 2005; Cosgrave, Bennie, & Kerslake, 2002; Kerslake, 1999, 2000, 2001). A further 23–25% of students were responding to the programme at the end of a school year and were carried over into the next year. According to the New Zealand monitoring reports, 84–87% of RR students each year have successfully completed the programme. Approximately 8% each year were referred for additional services, while 5–6% moved schools and did not complete the programme. Another 1–2% of students were unable to continue because of other factors such as the RR teacher moving and not being replaced (Anand & Bennie, 2004, 2005; Cosgrave et al., 2002; Kerslake, 1999, 2000, 2001).

Recent reports suggest that students who enter the programme with the lowest literacy levels may not be as likely to benefit from the programme as students with higher entry levels. RR monitoring reports for 2002 and 2003 provide data on students' entry levels on three measures used in the Observation Survey of Early Literacy Achievement (Clay, 1993a): RR text level, the Burt Word Reading Test (Gilmore, Croft, & Reid, 1981), and writing vocabulary and student outcomes (Anand & Bennie, 2004, 2005). Anand and Bennie stated that a student's initial entry-level results are "suggestive, but not predictive of, the likelihood of the student requiring referral" (2005, p. 14). They reported that the "average reading scores for the students, who eventually required referrals, were typically half the scores of students who successfully completed RR" (Anand & Bennie, 2004, p. 14). In addition, the students who were referred spent much longer in the programme, averaging 87 sessions in 2002 and 90 sessions in the programme in 2003 (Anand & Bennie, 2005). In 2002 this was 11 sessions more than those who were discontinued, and 13 sessions more in 2003. Students who were referred in 2002 typically made 64% of the gain of the successful students (Anand & Bennie, 2004) and those referred in 2003 averaged 66% of the gain of their discontinued peers (Anand & Bennie, 2005).

Reading Recovery in the United States

RR was introduced in the U.S. in 1984 in Columbus, Ohio (Lyons & Beaver, 1995). The programme spread rapidly and, by 1994, 47 states were implementing the RR programme (Lyons & Beaver). RR has been extremely successful in implementation in the U.S. and, by 2000, it was described as the "most widespread teacher-implemented, one-to-one intervention currently in use in schools in the U.S." (Elbaum, Vaughn, Hughes, & Moody, 2000, p. 606).

Despite the fact that evaluation data have been gathered for all students who have participated in the programme in the U.S., there is controversy about the interpretation of the data and the validity of the findings (Schwartz, 2005). Most "in-house" evaluation reports focus on the percentage of students who completed a full programme and who were discontinued. This percentage appears to be consistently around 80%, with Lyons (2003) reporting a discontinuation rate of 81% of students nationwide from 1984 to 1997 and, more recently, Gomez-Bellenge, Rodgers, and Schulz (2005) indicating that the national discontinuation rate for 2003 to 2004 was 76%.

Other evaluations focus on the percentage of students who started the programme rather than those who were discontinued. Lyons (2003), in reporting on trends shown in data collected over a decade, summarised RR's success as follows: "if we consider all students served, even for one day, 60% met the stringent exit criteria for success" (p. 218). Gomez-Bellenge et al. (2005) reported that, in 2003/04, 59% of all students served in RR programmes in the U.S. were discontinued, while another 18% of students who completed a full programme were recommended for additional services. These percentages are in line with Shanahan and Barr's (1995) calculation of 62% who were successfully discontinued from all children served in the U.S. at

that time. Esposito (2004) reported on an internal evaluation of RR implementation over a 15-year period in the Madison, Wisconsin school district. Over this period, only 57% of students in the programme were successfully discontinued. In 2003/04, 48% of the students (159 out of 305) in the programme were successfully discontinued. Analysis of the results of over 600 first-grade students in Durham, North Carolina, from the 1998/99 school year indicated that 40% did not successfully complete the programme (Haenn, 2000). Overall 43.3% of students were successfully discontinued, 30% were referred, 23% had incomplete programmes but were likely to complete the programme, 3.3% had moved, and 1% did not complete for other reasons.

Grossen, Coulter, and Ruggles (1997) suggested that success rates do not take into account the number of children who were eligible but were never served. For example, in an evaluation in Ohio, it was estimated that this group represented 19% of the students who were eligible for RR. Grossen et al. then calculated that "if the 19% who were never served are included, the success rate drops to 51%" (1997, p. 13).

Reading Recovery in England

RR began in England in 1990 with one tutor working in schools in Surrey (Hobsbaum, 1997). In 1992 a national pilot project was funded until 1995. The programme expanded quickly with the establishment of a national coordination network to monitor and coordinate the programme (Hobsbaum). Data from students in the programme in 1994 indicated that the programme was successful in raising children's levels of literacy, although only 70% of students who received a full programme were successfully discontinued, with 30% of students being referred for additional support (Hobsbaum). However, when calculated as a percentage of all students in the programme (i.e., including students who left the school or had incomplete programmes), the success rate was only 47%.

Reading Recovery in Australia

RR was introduced into the state of Victoria in 1984, implemented widely across that state, and has been introduced into all other states in Australia. The extent of coverage of the programme has varied from state to state, with the majority of research reports and data about the programme emanating from the states of Victoria and NSW.

In 2003, over 8,000 students (18.7% of the Year 1 cohort) participated in RR programmes in Victorian state schools (Victorian Department of Education and Training, 2003). The state has made a significant government commitment to RR since 1999, with over \$145 million allocated for the programme from 1999 to 2003 (Auditor General, Victoria, 2003). System data for 2003 indicate that 90.2% of students who completed the programme were successfully discontinued, 6.4% were recommended for further support, and 3.4% had an incomplete series of lessons (Victorian Department of Education and Training).

The Auditor General, Victoria (2003) investigated the effectiveness of RR in state schools in Victoria, tracking 42% and 50% of the students in the programme in 1999 and 2000 through to state testing in their respective Year 3. The results showed that neither group of students achieved as well as students who had not participated in the RR programme, even though programme-specific data indicated that the students did have improved reading proficiency (Auditor General, Victoria). A separate evaluation of literacy programmes in Catholic schools in Victoria found that RR students made gains similar to their peers while they were in the programme, but that these gains disappeared within 12 months (Ainley, Fleming, & McGregor, 2002).

In 1990, the NSW state education system officially began a RR pilot programme. The RR programme in NSW was subsequently extended with the commencement of a tutor training programme and with the allocation of greater levels of state funding. By the end of 2003 more than 2,100 teachers had undertaken RR teacher training, and there were 28 tutors working in the system (NSW Department of Education and Training, 2003). In 2004, RR was used in 837 schools by 920 teachers at a cost of AUD \$29.2 million (NSW Department of Education and Training, 2004).

The NSW Department of Education and Training has published data about the outcomes of the state's RR programme in government schools in its Annual Reports. The report for 2003 indicated that 70% of the RR students who were discontinued in 2001 "demonstrated expected levels of achievement or higher (were in Bands 2 or higher)" in the literacy component of the Year 3 Basic Skills Test (BST) in 2003 (NSW Department of Education and Training, 2003, p. 9). Results of this state-wide assessment are reported in five bands of achievement, with Band 1 being the lowest. The percentage of RR students achieving at Band 2 or higher in the 2003 assessment was slightly lower than reported for previous years; 76% in 1999, 74% in 2000, 72% in 2001, and 75% in 2002. On the surface this appears to indicate that the programme is effective, but closer inspection and analysis shows that the programme is not as effective in NSW schools as it is claimed.

When compared with state averages for all students, it is obvious that the majority of the ex-RR students did not achieve state averages or higher as a group. The average levels for all students who sat the Year 3 BST in 2003 was at least Band 3, as only 28% of students were in Bands 1 and 2 in 2003, 28% in Band 3, and 44% in Bands 4 and 5 (NSW Department of Education and Training, 2003). In comparison, 64% of ex-RR students were in Bands 1 and 2 in 2003. The data for 1999–2003 are relatively consistent in showing that approximately only one in three ex-RR students achieved state-average levels for literacy in Year 3. Thus it cannot be claimed that RR has achieved its goal of returning students to the average levels of the class and providing them with a self-improving system that will maintain their accelerated rate of progress.

In its 2004 Annual Report, the NSW Department of Education and Training (2004) used a slightly different type of reporting in which it presented the percentage of successful RR students who reached the minimum literacy standard; that is, performance at Band 2 level, or higher, in the Year 3 BST, and at Band 3 or higher

in the Year 5 BST. For Year 3 students this percentage was 77.4%, in comparison with 89.2% of the full Year 3 cohort who achieved Band 2 or higher. The percentage of Year 5 students was 84.7%, in comparison with 93.1% of the full cohort. These figures indicate that some successful RR students do not retain gains made in the programme.

The NSW Department of Education and Training's (2003) Annual Report also provides information about maintenance levels of ex-RR students in Year 5. It reported that 86% of Year 5 students who were discontinued from the RR programme in 1999 "demonstrated expected levels of achievement or higher (Band 3 or higher) in the BST literacy test in 2003 compared to 88% in 2002" (NSW Department of Education and Training, 2003, p. 9). The results of this assessment are reported in six bands, with Band 1 being the lowest. The average level for all students who sat the Year 5 BST in 2003, however, would be above Band 3 as there were only 19% of students in Bands 1, 2, and 3. In comparison, 43% of ex-RR students were in Bands 1, 2, and 3 in the Year 5 tests. Again, it has been demonstrated that the majority of ex-RR students did not retain gains made in the programme.

Looking Back Across Systems

From RR's data and system data from implementation in four countries it appears that RR has spread rapidly in school systems and has short-term effects for the many students who are successfully discontinued. System data that look at long-term effects, however, indicate that short-term gains dissipate over time.

In the systems investigated in the present review, the percentage of RR students who complete a full programme who are discontinued is around 70–80%—except for New Zealand and the state of Victoria in Australia, where the percentage (from in-house evaluations) is reported to be between 84% and 90%. When success is calculated as the percentage of students who start the programme who are discontinued, all systems except New Zealand have success rates between 48% and 65%, while in New Zealand this percentage is reported to be 84–87%.

Comparisons of levels of student change across systems are difficult to make because the programme operates slightly differently and so data are not always comparable. For example, students with incomplete programmes at the end of Year 1 usually do not continue the programme the next year in school systems in the United States, yet in New Zealand most of these students will have an opportunity to complete the programme in the following year. This has a significant influence on the percentage of students who begin the programme and who are successfully discontinued. It is also difficult to identify the group of "carried-over" students in some evaluation reports. In some reports it appears that these students are counted in two years' programme numbers if they are carried over to the next year. This may mean that fewer students than those reported are actually served.

In-house evaluations provide data on the students who complete the programme and, in some instances, ignore the students who are withdrawn from the programme

or have incomplete programmes (Landis, 1997; Shanahan & Barr, 1995). While it is difficult to factor in the number of students who do not complete the programme, it is misleading to ignore this group (Grossen et al., 1997; Snow et al., 1998). The possible effect of this is that the effectiveness of the programme is overstated or inflated (Grossen et al.).

Experimental Research

It has been suggested that the evidence about RR is unclear because of flaws in research methodology and lack of independence of those gathering or analysing data (Grossen et al., 1997; Hiebert, 1994; Shanahan & Barr, 1995). In the past 15 years there have been only a limited number of well-designed experimental studies featuring RR, and these provide evidence about effectiveness, programme delivery, and the teaching programme for children.

An evaluation of the RR pilot programme in NSW by Center et al. (1995) was one of very few research studies on RR with randomised experimental and control groups. It also included a comparison group of students of similar reading ability attending schools where RR was not implemented. The authors found, after 15 weeks of instruction, that RR students performed better than control students on text-reading and word-reading tests but there were no significant differences on two out of three tests that measured metalinguistic skills (Center et al.). The effect sizes were smaller (on all tests but one) when short-term maintenance testing occurred 3 months after the post-test. Student data indicated that 65% of the students in the RR group had been successfully “recovered” to average class levels (Center et al., 1995, p. 260). This is in line with many other evaluation studies.

Of interest were the data on medium-term testing of students (1 year after the post-test), which showed that 29% of students in the comparison group and 30% of the control group (after adjustments for loss of students to the experimental group) had recovered without RR (Center et al., 1995). If these students recovered without intervention, it may reasonably be assumed that the same percentage of RR students would have also done so. Center et al. concluded that RR was, therefore, effective for only one in three students undertaking the programme.

An unpublished report also indicated that students in the first and second intakes into RR in Year 1 had quite different outcomes (Wheldall, Center, & Freeman, 1992). Sixty-two per cent of first intake students were discontinued, in comparison with only 36% of second-intake students. However, as 28% of comparison students had recovered without intensive instruction, it was suggested that RR was only effective for 34% of the first-intake students and a mere 8% of second-intake students.

In a recent study in the U.S., Schwartz (2005) investigated the effectiveness of RR and compared the results of students in the first intake of students in Grade 1 with students in the second intake. In this study, 37 RR teachers nominated two at-risk students each. These students were then randomly assigned to either a first-intake or a second-intake RR group. Comparisons were then made during the year between students with and without intervention and with a low-average and a

high-average student from the same classroom as the first-intake student and the second-intake student. Students were assessed at the beginning of the year, at the transition between intakes, and at the end of the year, using the Observation Survey of Early Literacy Achievement (Clay, 1993a). They were also tested during the transition period and at the end of the year on phonemic awareness tests, an oral reading test, and a comprehension test. Overall, 65% of students from both intake groups were discontinued, 16% were recommended for further services, and 16% had incomplete programmes, one student moved, and another withdrew.

The results show that the at-risk students who received RR during the first intake performed better than students who did not receive RR during this period. This was most obvious on the tests used in the Observation Survey of Early Literacy Achievement (Clay, 1993a) and the oral reading test, all of which showed large effect sizes. There were no significant differences among groups on the measures of phonemic awareness or comprehension at the transition period. Schwartz (2005), using a similar process to that in the Center et al. (1995) study, reported that 14% of second intake students had made reasonable progress in reading by the transition point without any intervention, indicating that RR may service some students who would have “recovered” anyway. The author also found that at-risk students who received RR, when compared with low-average classroom peers, closed the achievement gap. It should be noted that, although some of the comparison groups of classroom peers entered RR during the second half of the year, their data were used in comparisons at the end of the year.

Pinnell et al. (1994) compared RR with three other literacy interventions for young at-risk students: Reading Success (a one-to-one tutorial programme based on RR), Direct Instruction (a one-to-one skills-based intervention), and a small reading and writing group that was taught by a trained RR teacher. The U.S. study had 403 subjects who were randomly assigned to treatment groups. Each treatment group had its own comparison group. RR was the only intervention that showed a “mean treatment effect” on two standardised measures and two RR measures, indicating that the effect does not just occur because it is a one-to-one tutoring programme. The authors looked at programme characteristics and surmised that the reasons for success in tutorial programmes are “individual instruction, instructional emphasis, and teacher professional development” (Pinnell et al., 1994, p. 36).

A quasi-experimental study by Quay, Steele, Johnson, and Hortman (2003) was conducted with 214 students in 34 schools in a U.S. school district in the first year of implementation of RR. It investigated the effects of RR delivered by teachers-in-training on student achievement and five personal and social aspects. Students were randomly allocated to classes and then one class in each school was randomly designated to be the class that received RR. Control subjects were drawn from another classroom. The authors found that the RR students taught by RR teachers-in-training were superior to the control students on standardised tests and all of the tasks on the Observation Survey of Early Literacy Achievement (Clay, 1993a) when tested at the end of first grade (Quay et al.).

Iverson and Tunmer (1993) conducted a study in the U.S. comparing RR with a modified RR instruction (one-to-one) and a standard intervention group. The modified RR group received a standard RR programme that included explicit instruction in letter-sound patterns instead of letter identification procedures. They used matched groups of 32 students who were pre-tested and post-tested on RR measures, a word recognition test, and tests of phoneme segmentation, phoneme deletion, and phonological recoding. The study showed that RR was “highly effective” (Iverson and Tunmer, 1993, p. 123) as students in both RR groups demonstrated that they were reading at the level of their peers on completion of their programme and at the end of the year. This, in itself, may have been the effect of individual instruction compared with group instruction. The modified RR students learned to read much more quickly than the RR students, taking an average of 41.75 lessons to reach discontinuation point, whereas the RR students took an average of 57.31 lessons.

Another experimental study investigating possible modifications to RR was conducted by Iverson, Tunmer, and Chapman (2005) in New Zealand. They compared the literacy outcomes of students who were instructed in a modified RR lesson with the outcomes of students receiving traditional one-to-one RR instruction. They found no significant differences between the two groups on any measures and between groups on the number of lessons, although the modified RR lessons took an average of 42 min compared with 32 min for RR instruction.

Multilevel modelling was used by Plewis (2000) to evaluate the effects of the RR programme in England. RR was compared with a control group with no intervention and a phonological training intervention based on the work of Bradley and Bryant (1985). Students in the phonological intervention were taught by trained tutors in 40 × 10-min sessions over a 7-month period. The six poorest readers were selected in 63 schools, 22 of which were RR schools. In the RR schools the bottom three or four readers entered RR while the remaining two or three formed a within-school control group. All six students in 18 control group schools were controls, and in the 23 schools with phonological training four of the six students were randomly assigned to the intervention group, with the remaining two comprising the control group. Follow-up measures were taken after 1 year, 2 years, and 4 years. Plewis found short-term effects for RR with student reading gains of between 3 and 11 months in reading age. However, gains were not retained in comparison with a control group after 3 years.

Reviews of Research

In the past 15 years there have been several research reviews that have investigated aspects of RR. The first of these by Wasik and Slavin (1993) reviewed research on RR as one of five tutoring programmes aimed at preventing early reading failure. At that time there were only two longitudinal studies that were suitable for inclusion. The authors found that students made initial gains in the implementation year with an effect size of more than +0.7 for both pilot and second cohorts. However, these effects washed out after 12 months and were further reduced at a 2-year follow-up. In addition they found that, while the students who were successful were reading at

average levels or above at follow-up testing, students who were not discontinued (27% of those tested at the 2-year follow-up) were still reading below the average level of their peers (Wasik & Slavin).

Elbaum et al. (2000) also conducted a quantitative review of interventions featuring tutoring programmes but focused solely on those that delivered one-to-one instruction. Of the 29 studies, 16 featured RR. They found that RR had an effect size of 0.66, significantly greater than that of other matched interventions. Following adjustments for teacher training and as a result of two studies comparing small-group instruction with one-to-one tutoring, the authors concluded that their meta-analysis “did not provide support for the superiority of RR over other one-to-one interventions” (Elbaum et al., 2000, p. 617).

A meta-analysis of 36 U.S. studies by D’Agostino and Murphy (2004) indicated that RR has positive effects. The studies were conducted between 1984 and 1995. Almost all were in-house reports and evaluations, with only one having an experimental design and very few having been published in peer-refereed journals. As these evaluations were not independent, the results of the meta-analysis should be interpreted cautiously. The authors found that RR had positive effects for both discontinued and not-discontinued students on assessments designed for the programme and on standardised measures. RR effects were greatest for discontinued students on programme measures, however.

Hiebert (1994) conducted a review of RR in the U.S. over the 10 years from 1984 to 1993 to determine its capacity to change “the literacy profiles of age cohorts” (p. 15) and its cost-effectiveness. Noting differences between the education system in New Zealand and those in the U.S., she stated that RR had not been adapted to American needs, the effectiveness of the programme could not be determined from the available evidence, and that further investigation was needed.

A comprehensive review of RR by Shanahan and Barr (1995) looked at student progress in literacy learning, retention of gains, effects on classroom instruction, professional development, cost-effectiveness, and research. In regard to the effect of RR on student learning, the authors found that “RR works, but not as well as its proponents have claimed” (Shanahan & Barr, 1995, p. 989). They surmised that it had “limited maintenance” (p. 990), with short-term gains decreasing in relation to the achievement of average students. Shanahan and Barr stated that, while recognising that there was a scarcity of research in this area, there was little evidence of RR having an impact on what happens in the classroom. They also indicated that RR was a relatively expensive intervention, costing around USD\$4,000 per student (equivalent to approximately AUD\$5,400) after factoring in savings from the reduced demand for later special education services and lowered retention rates.

Research about Long-term Effects

While research studies consistently show that discontinued students make short-term gains, there are indications that there may be a “wash-out” effect over time (Wasik & Slavin, 1993). In one study, gains from RR were generally not retained on

state-wide assessments in third grade (Haenn, 2000). Chapman, Tunmer, and Prochnow (2001), in a longitudinal study that concluded RR did not overcome deficits in phonological skills, found that students who had been discontinued from the RR programme, when tested 1 year after completing the programme, had reading achievement levels approximately 12 months behind their chronological age.

Hiebert (1994) searched for evidence of long-term effects and noted that the only longitudinal study that was available was poorly designed and did not have a significant finding. She analysed data from this study and found that only 5.5% of ex-RR students would have had improved reading achievement at the end of Grade 4 as a result of the programme. Pinnell, Lyons, and Jones (1996) explained this as being influenced by factors such as problems with instruction by teachers in later grades, policies on retention of students, and students' individual circumstances.

On a more positive note, a study by Brown, Denton, Kelly, and Neal (1999) followed up over 600 RR students in Grade 5 in California and found that 75% of students with complete programmes had average or above-average scores on standardised tests. A study into professional development in Victoria in Australia by Rowe (1997) found that RR students "benefited notably by participation" (p. 76) and that early literacy gains were maintained in Grades 5 and 6. These findings are encouraging.

Some researchers have investigated the effects of RR on the system. Lyons and Beaver (1995) conducted a longitudinal study in Ohio and found that RR significantly lowered the number of students who were placed in learning disabilities programmes in the second year of schooling, and it also reduced the percentage of students who were retained. In this study, the number of students identified as learning disabilities in one school district was reduced from 36% in 1986/87 to 6% in 1990/91. The authors also noted that the percentage of students who were retained in their grades was reduced by two-thirds during this time period. Other system effects were noted by Gredler (2001), who reported a reduction of 69% in the number of students enrolled in transition programmes in New Hampshire that cater for students identified as being "developmentally immature" (p. 24).

Looking Back at What Reading Recovery Has Done Well and Not So Well

Looking at the rapid growth of RR in a number of education systems, it is obvious that the programme has established an enviable reputation as an intervention. Over the past 15 years, however, a number of concerns have been voiced about aspects of the programme (Hiebert, 1994; Shanahan & Barr, 1995), indicating that, while there are many things that RR has done well, there are things that RR has not done well.

What Reading Recovery has Done Well

1. RR has been shown to be an effective intervention in the short term for many students (Center et al., 1995; Haenn, 2000). However, there are disparities

between the percentages of successful students reported through “in-house” data collections and those through independent evaluations, with the latter mainly reporting success rates of approximately 60% of all students who enter the programme.

2. RR is designed for implementation at what is seen by Clay (1991) as an optimal time for young learners. While the notion of a preventative programme for young literacy learners is supported by research that indicates early intervention to prevent literacy failure is crucial (Juel, 1988; Stanovich, 2000), it is possible that the notion of an optimal time could be narrowed down within Year 1 or could be during the kindergarten year. In support of this view, one study indicated that students who entered the programme in the first intake in Grade 1 had markedly better outcomes than students who entered the programme in the second intake (Wheldall et al., 1992).
3. RR includes many components of successful early reading instruction (Hiebert, 1994). These include high expectations, time spent reading and writing, rereading of texts, setting clear goals, learning about letter-sound relationships, making time for observation of students’ reading, deliberate teaching, phonemic awareness, and professional development that focuses on effective instruction (Hiebert).
4. RR has an effective implementation process, demonstrated by the fact that it has spread widely and relatively rapidly through a number of education systems. In most cases, Marie Clay and personnel from New Zealand have played a key role in establishing the programme within a system. This has been a means of maintaining quality control in tutor (teacher leader) and teacher training, in provision of information for administrators, schools, and teachers, and advocacy for funding and administrative support. In each system investigated, RR has been implemented according to the principles of teacher change, school organisational change, and political and funding changes set out in Clay (1987), although with some modifications (e.g., the stage of schooling when a student enters the programme).
5. RR has been particularly successful in gaining political support in many education systems. This is critical for allocation of funds for teacher training, tutor (teacher leader) training, administrative functions, and teacher salaries. Its success in gaining political support in NSW is, for example, demonstrated by a dispute between the two main political parties prior to the 2003 state election, with both claiming that they were responsible for introducing RR in state schools (Dempster, 2003). In the U.S., however, RR is under threat of losing funding for implementation in a number of states (RR Council of North America, 2006). Consequently, the RR Council of North America has responded to this threat or “attack” (2006, p. 1) by providing a rebuttal of a number of criticisms of the programme.
6. Teacher selection, training, and on-going professional development have been praised (Shanahan & Barr, 1995). RR teachers are trained to implement Clay’s theories and “approach reading instruction with a deep and principled

understanding of the reading processes and its implications for instruction” (Snow et al., 1998, p. 258). The role of RR teacher is usually sought after and has special status among staff members in a school as a teacher with a high level of knowledge. RR teachers generally have a great sense of pride in their role and a sense of belonging to a special group.

7. RR collects individual student data from all students in the programme for analysis. This provides an extensive amount of data about student progress that can be used by programme and system administrators to monitor RR’s effectiveness.
8. RR has established a reputation as an effective intervention among educators and administrators. It has achieved this by publication and dissemination of the results of Clay’s field trials, site reports, evaluations, and research studies (Clay, 1987). Clay has also encouraged tutors (teacher leaders) and administrators to counter criticisms or misunderstandings.

What Reading Recovery has Not Done Well

1. While the programme has shown that it works for many students, it has not demonstrated that it works for the students who are most at-risk of failing to learn to read (Baker et al., 2002). Haenn (2000), in reviewing the pre-test and post-test results of 610 students in Durham, North Carolina, found that the students who were lowest on most measures of the Observation Survey of Early Literacy Achievement were least likely to succeed on the RR programme.
2. RR has not been shown to facilitate reading development in students with poor phonemic awareness (Center et al., 1995; Chapman et al., 2001). Students who enter the programme typically have poor phonological processing skills and those with the lowest scores have been found to be least likely to benefit from the programme. In addition, Chapman et al. found, when RR students were tested after the intervention, that RR “did not eliminate deficiencies in phonological processing” (2001, p. 158) and that problems in this key component of early reading impacted on students’ progress in reading in both the short term and the long term.
3. RR does not reflect recent research findings about the crucial components of an early literacy programme (Groff, n.d.; Tunmer & Chapman, 2003). Moats (2000) suggests that RR is “a whole-language incarnation” that uses whole-language practices such as teaching students to predict unknown words from context and initial letter cues, incidental phonics instruction, use of running records, and decoding by analogy within a structured lesson format. RR has been the subject of criticism because the theoretical principles and teaching procedures suggest that context is more important in predicting up-coming words than graphophonic cues (Tunmer & Chapman), even though research very strongly suggests otherwise (Stanovich, 2000).
4. Although it has been implemented for 15–20 years in some education systems, RR has not demonstrated that it has dramatically reduced literacy failure within education systems (Auditor General, Victoria, 2003; Shanahan & Barr, 1995;

Welna, 1999). In some studies, research has shown that short-term gains are often not retained and that these have all but disappeared by Year 3 (Hempenstall, 1999; Shanahan & Barr), Year 4 (Hiebert, 1994), or Year 5 (Snow et al., 1998). While there have been some reports of maintenance of gains, lowered retention rates, and lowered referrals to special education, there is no evidence of a dramatic reduction in literacy failure in education systems since RR was introduced. This may be caused by limited budgets in education systems resulting in resources being spread so thinly that not all struggling Year 1 readers can access the programme, or it may be an indication that RR has limited or differential long-term effects.

5. RR has relatively high costs, and doubts have been cast on its cost-effectiveness within a system (Hiebert, 1994; Shanahan & Barr, 1995). Groff (n.d.) suggests that those who promote RR typically downplay its cost, not taking into account the cost of teacher training, additional benefits for teachers, resources, salaries, and travel expenses for RR programme administrators. Calculations vary, but have been reported to be as high as AUD\$9,088 per student in New South Wales in 2003 (NSW Department of Education and Training, 2003) and as low as USD\$1,708 (equivalent to AUD\$2,350 in 2006) per student in Lancaster City School District in Ohio (Lyons & Beaver, 1995). Groff (n.d.) suggests that RR is not cost-effective in systems where the literacy programme has a different theoretical basis and, therefore, is not complementary.
6. In some systems RR does not target students well. The practice of targeting the lowest 20% of students in a school results in some students in higher-achieving schools receiving services, while Grade 1 students with literacy difficulties in lower-achieving schools may not be able to access the programme. Rather than targeting the lowest 20% in a school, funding would be more equitably directed to those students who are performing at the 20th percentile and below. System differences at time of entry to the programme also influence the identification of students for entry to RR. For example, Shanahan and Barr (1995) note that at age 6, while students in New Zealand have completed a year of schooling before they enter RR, students in the U.S. are only starting first grade and prior to this have often had little schooling. As RR programmes in the U.S. maintain the same age of entry to the programme, it is likely that many students who would not need intervention are placed in the programme (Shanahan & Barr).
7. Despite the large amount of data collected within the programme, RR has a relatively weak research base. There are a limited number of true experimental studies about RR's efficacy that feature randomly allocated groups (Center et al., 1995; Snow et al., 1998). These are the most effective designs in demonstrating causality and controlling threats to internal validity, and are described by Seethaler and Fuchs (2005) as providing the "highest level of proof of efficacy of a programme" (p. 99). Center et al., in one of the few studies that used an experimental design, demonstrated that a significant percentage (up to 30%) of those RR students who were successfully discontinued may have recovered without RR.

8. RR does not use independent measures of reading development, and, thus, reports of student progress need to be interpreted cautiously (Tunmer & Chapman, 2003). In many studies, student reading development is measured by using instruments taught within the programme, so it would be expected that students would progress on these measures (Baker et al., 2002; Grossen et al., 1997). Although several experimental studies rely on external reading measures, standardised tests are not often used as measures of progress within the programme (Snow et al., 1998). In some evaluation studies, reading achievement is reported in relation to movement through RR text levels. This has been seen as problematic because the text levels were developed in quite an arbitrary manner and there are not equal intervals between each book level (Baker et al.; Grossen et al.).

Looking to the Future

After more than two decades of implementation of RR in many systems, there has been sufficient time and opportunity for system administrators and governing bodies to observe and examine the effects on their education systems. It is timely, then, to consider the long-term future of the programme and to examine those aspects of the teaching programme that it does not do well, to consider whether the programme can be improved or whether alternatives are viable. These aspects include making the programme more effective for the students who are most at-risk, improving success rates overall, retention of gains, and lowering costs. Issues such as the selective reporting of programme data and the need for experimental research are separate concerns.

Some authors suggest that changes to the existing RR programme would make it more effective, in particular for those students who are most at-risk. A recurring recommendation from research and reviews is that RR developers make changes to the teaching of phonological processing skills. For example, in 1999 a group of literacy experts was asked by the New Zealand Ministry of Education to report on literacy instruction in New Zealand. One of its recommendations was that RR “place greater emphasis on explicit instruction in phonological awareness and the use of spelling-to-sound patterns in identifying unfamiliar words in text” (Ministry of Education, 1999, p. 6). This recommendation is in line with information about the key components of effective early literacy instruction that has emanated from major reviews of the teaching of early literacy in the U.S. (National Institute of Child Health and Human Development, 2000), Australia (Department of Education Science and Training, 2005), and England (Rose, 2006).

RR has an in-built process in which some changes can be made to the programme and disseminated to RR teachers through teacher development and professional support processes. This enables the designers of the programme to inform teachers and incorporate new information from research into the teaching programme. However, due to the quality control processes in place, any changes need to be sanctioned by programme designers and developers.

In recent years, changes have been made to the teaching programme to incorporate research findings about phonological processing (Schwartz, 2005). RR personnel state that they “give specific and explicit attention to letters, sounds and words” (RR Council of North America, 2006, p. 20) but this does not necessarily mean that these aspects are taught either explicitly or systematically. A systematic approach to teaching phonics has been found to be the most effective approach for teaching at-risk readers (Ehri, Nunes, Stahl, & Willows, 2001), yet is at odds with some of the principles of RR described by Cox and Hopkins (2006). The use of a systematic approach to teaching would require a change in the underpinning principles of RR away from its constructivist approach to a direct teaching approach. Theoretically, while minor variations can be made to the content of the programme by the developers, it is extremely unlikely that changes of this nature would be considered.

For RR to be maximally effective in a school system, it has been suggested that it needs school contexts that are conducive to, or compatible with, the programme. Some recommendations include that a “supportive, substantive, and coherent plan of reading/literature instruction” exists in all grades (Landis, 1997, p. 3) or that RR be implemented with “depth” within a school (Munn & Ellis, 2005, p. 357). Munn and Ellis found that students in schools where the RR programme was not “deeply implemented” (2005, p. 355) spent an average of 4 weeks longer in the programme than students in schools where the programme had “depth” (p. 355).

Center, Freeman, and Robertson (2001), however, found that RR students did better if their regular classroom was “code-orientated” rather than “meaning-orientated” (p. 224). This finding may be related to findings about the efficacy of an explicit teaching approach in comparison with an implicit approach, to the findings about the key role that phonological processing plays in early reading success, or to both factors. Further research is warranted to investigate the most effective combination of intervention and classroom contexts for students who are experiencing difficulties in early literacy learning.

It has been suggested that the cost of one-to-one interventions such as RR could be reduced significantly if they were to be delivered in pairs (Fawcett, Nicolson, Moss, Nicolson, & Reason, 2001; Iverson et al., 2005) or in small groups (Elbaum et al., 2000). The study by Fawcett et al. suggested that an intervention could be delivered to pairs with comparable results to RR at about 20% of the cost. On the surface, the notion of providing tutoring to pairs or small groups may be appealing, but would mean that it would make a significant change to the way RR teachers plan and deliver an individual programme for each child. To keep the RR programme intact, Clay has been very firm in insisting that RR cannot be tampered with and has trademarked the name (RR Council of North America, 2004b). Therefore, it is highly unlikely that a variation of RR would be (legally) implemented.

One of the key factors in reducing the cost of an intensive intervention such as RR is accurate identification of target students. It is obvious that, if students do not acquire early reading skills for whatever reason, they will have low scores on the assessment battery and are likely to be targeted for RR intervention. One possible reason could be that students are poorly taught in the first year at school, and a

feasible solution is that instruction could be improved in the first year of schooling (Hempenstall, 1999).

It is possible that there is over-identification of students who need intervention with a significant accompanying cost to the funding provider. Wheldall, Center, and Freeman (1993) found that a large percentage of students who were selected for RR would have "recovered" without intervention. A recent study by McCusker and Munro, as yet unpublished, supports this finding. McCusker and Munro tracked 700 students in Catholic schools in Victoria, comparing at-risk students who accessed the RR programme with students who were also at-risk but did not access the programme and with students who made good progress. They found little difference in achievement between the two at-risk groups and also found that some of the students who were in the at-risk group at the time of identification achieved scores in the average range after 12 months, even though they did not receive any intervention (H. McCusker, personal communication, 13 January 2006). Further research in this area may assist in developing a more accurate identification of students for RR with resultant cost savings.

Given that research on the long-term effects of the programme is equivocal and the effects that have been found are neither strong nor consistent, educational planners need to be aware that the intervention is unlikely to dramatically reduce the number of students who will require special education services in later years. The notion of a large decrease in illiteracy rates is a compelling argument for a system to implement RR in the first place, but it appears that, instead of counting on long-term savings in future years, education systems need to ensure that there is a continuing budget for special education services throughout all of the years of schooling.

If RR is to maintain its reputation as an effective programme it needs to demonstrate to the general research community that it has a strong research base. To date, despite the collection of a great deal of data, there is limited experimental research to support the efficacy of the programme. Well-designed experimental research that uses randomly assigned intervention groups is needed to investigate aspects that have equivocal findings. In trying to plan research with randomly assigned groups, one of the issues is that a RR teacher with a normal load identifies and teaches the four lowest-performing students in a school and that, in this situation, random assignment cannot occur. In addition, the conduct of research or reviews is made difficult by the reporting of student progress in terms of the programme's own measures. As a result, much of the "research base" for RR is not independent. If RR is to prove its success to educators, the use of independent measures would provide a firmer base of evidence. Looking to the future, research would be facilitated if RR administrators and independent researchers collaborate on ways to undertake quality studies without compromising the intervention.

In planning for the future, it is also important to determine whether there are alternatives to RR that are research-based. If there are not, interventions that are based on research should be developed, and trials and comparison studies should be conducted. Some educators suggest alternatives to RR (Haenn, 2000; Nicolson,

Fawcett, Moss, Nicolson, & Reason, 1999) but there appear to be very few that have been thoroughly researched and have strong evidence to support their use. One example is the programme Early Steps, a first-grade intervention programme like RR, with the addition of direct and systematic instruction in orthographic patterns (Morris, Tyner, & Perney, 2000). It was found to be highly effective, especially for those students who “were most at risk” (Morris et al., 2000, p. 687). Another intervention that was trialled with groups of four students in their first year at school in the United Kingdom featured the teaching of phonics skills and word-building, and broader reading, using a computer-assisted programme (Nicolson et al.). The authors found that the intervention group had significantly greater gains than a control group, although 25% of the group remained at-risk readers. The mean effect size of 1.71 for reading age was reported to be comparable with RR at 10% of the cost of RR. These studies are of small scale and would require much more evidence of success in a number of settings before recommendations could be made for their use.

Looking back at research over the past 15 years, RR appears to be one of the most widely, if not *the* most widely, evaluated interventions in education. While there have been criticisms of its effectiveness, especially for students who are most at-risk, there appears to be no better alternative that is widely available, at the present time. While it can be argued that it is not highly effective, it is effective for many students and appears to be no worse than other early literacy interventions.

In the future, new interventions could be designed, trialled, and researched. Many of the teaching strategies within RR are commonly used by regular classroom teachers in classrooms and are features of other interventions. Some, or all of these, could be incorporated into a tutoring programme that also features content and strategies drawn from recent research into reading acquisition. Crucial components of early reading programmes, including phonemic awareness, phonics, the reading of a large amount of text, and linking reading and writing (Snow et al., 1998), would ideally form the basis for the content of a new intervention. An alternative early intervention programme that is based on current research about what to teach and how the components are most effectively taught could be delivered in small groups, thereby making it more cost-effective.

Conclusion

This review, in looking back at over nearly two decades of implementation of the RR programme, has documented the extraordinary growth of the intervention across a number of education systems and has looked for the reasons for its success. RR has provided an excellent model in demonstrating how to plan, promote, and implement an intervention across an education system and how to design a professional development programme. These have been important elements in RR’s rapid spread across education systems.

Although RR has established a reputation as being a remarkably successful intervention (Johnston & Allington, 1991), research, however, indicates that it has

not delivered all that it promised to deliver: long-term change for students and a significant reduction in demand for special education services in later years. Evidence indicates that RR is beneficial for those students who are discontinued but that it is less beneficial for students who have incomplete programmes, are withdrawn, or are referred to special education. In fact, the success of the programme appears to be inversely related to the severity of the reading problem. A student who enters the programme with relatively high scores on Clay's (1993a) Observation Survey of Early Literacy Achievement is likely to be a success, while a student with a severe problem is unlikely to be a success.

The data regarding long-term system effects are equivocal and there is little evidence of the cost of special education services being reduced in later years. Given that the programme has been implemented widely in some education systems for a number of years, it is reasonable to expect that special education budgets in these systems would reflect lowered demands for services beyond Year 1 by now.

Although there are some alternative preventative programmes, there are none that have been implemented on the scale of RR or that, at the present time, have the infrastructure or research support to replace RR as a system intervention. It is even possible that the apparent success of RR may have inhibited the development of alternative early literacy intervention programmes. RR has shown that it is good, but it could be even better. In the future, educational researchers may develop an alternative intervention that is more effective. If so, it is recommended that they take lessons from RR about how to go about developing and implementing an educational intervention.

References

- Ainley, J., Fleming, M., & McGregor, M. (2002). *Three years on: Literacy advance in the early and middle primary years*. Melbourne, Australia: Catholic Education Commission of Victoria. Retrieved April 5, 2006, from <http://www.dest.gov.au/archive/schools/publications/2002/threeyearson/larp.pdf>
- Anand, V., & Bennie, N. (2004). *Annual monitoring of Reading Recovery: The data for 2002*. Wellington, New Zealand: Ministry of Education.
- Anand, V., & Bennie, N. (2005). *Annual monitoring of Reading Recovery: The data for 2003*. Wellington, New Zealand: Ministry of Education.
- Askew, B. J., & Frasier, D. F. (1997). Sustained effects of Reading Recovery intervention on the cognitive behaviours of second grade children and their perceptions of their teachers. In S. L. Swartz & A. F. Klein (Eds.), *Research in Reading Recovery* (pp. 18–38). Portsmouth, NH: Heinemann.
- Askew, B. J., Kaye, E., Frasier, D., Mobasher, M., Anderson, N., & Rodriguez, Y. (2003). Making a case for prevention in education. In S. Forbes & C. Briggs (Eds.), *Research in Reading Recovery* (Vol. 2, pp. 133–158). Portsmouth, NH: Heinemann.
- Auditor General, Victoria (2003). *Improving literacy standards in government schools*. Retrieved May 26, 2005, from http://www.audit.vic.gov.au/reports_par/agp9000.html
- Baker, S., Berninger, V. W., Bruck, M., Chapman, J., Eden, G., Elbaum, B., et al. (2002, May 20). *Evidence-based research on Reading Recovery*. Letter to members of the US Congress signed by 31 reading researchers. Retrieved March 15, 2005, from <http://www.education-news.org/Curriculum/Reading/ReadingRecoveryisnotsuccessful.htm>

- Bradley, L., & Bryant, P. (1985). *Rhyme and reason in reading and spelling*. Ann Arbor, MI: University of Michigan Press.
- Brown, W., Denton, E., Kelly, P., & Neal, J. (1999). Reading Recovery effectiveness: A five-year success story in San Luis Coastal Unified School District. *Educational Research Service Spectrum: Journal of School Research and Information*, 17, 3–12.
- Center, Y., Freeman, L., & Robertson, G. (2001). The relative effect of a code-oriented and a meaning-oriented early literacy program on regular and low progress students in Year 1 classrooms which implement Reading Recovery. *International Journal of Disability, Development and Education*, 48, 207–232.
- Center, Y., Wheldall, K., & Freeman, L. (1992). Evaluating the effectiveness of Reading Recovery: A critique. *Educational Psychology*, 12, 263–274.
- Center, Y., Wheldall, K., Freeman, L., Outhred, L., & McNaught, M. (1995). An evaluation of Reading Recovery. *Reading Research Quarterly*, 30, 240–263.
- Chapman, J. W., Tunmer, W. E., & Prochnow, J. E. (2001). Does success in the Reading Recovery program depend on developing proficiency in phonological-processing skills? A longitudinal study in a whole language instructional context. *Scientific Studies of Reading*, 5, 141–176.
- Clay, M. M. (1987). Implementing Reading Recovery: Systemic adaptations to an educational innovation. *New Zealand Journal of Educational Studies*, 22, 35–58.
- Clay, M. M. (1991). *The early detection of reading difficulties* (3rd ed.). Auckland, New Zealand: Heinemann Education.
- Clay, M. M. (1993a). *An observation survey of early literacy achievement*. Auckland, New Zealand: Heinemann Education.
- Clay, M. M. (1993b). *Reading Recovery: A guidebook for teachers in training*. Auckland, New Zealand: Heinemann Education.
- Clay, M. M. (1998, February). *The challenge of literacy improvement*. Paper presented at the NSW Department of Education and Training State Literacy Strategy Conference, Sydney, Australia.
- Cosgrave, R., Bennie, N., & Kerslake, J. (2002). *Annual monitoring of Reading Recovery: The data for 2001*. Wellington, New Zealand: Ministry of Education.
- Cox, B. E., & Hopkins, C. J. (2006). Building on theoretical principles gleaned from Reading Recovery to inform classroom practice. *Reading Research Quarterly*, 41, 254–267.
- D'Agostino, J. V., & Murphy, J. A. (2004). A meta-analysis of Reading Recovery in United States schools. *Educational Evaluation and Policy Analysis*, 26, 23–38.
- Dempster, Q. (2003, 21 March). Bob Carr faces the stateline panel, In *Stateline*. Sydney, Australia: Australian Broadcasting Commission. Retrieved April 13, 2005, from www.abc.net.au/stateline/nsw/content/2003/s813125.htm
- Department of Education, Science and Training. (2005). *Teaching reading: Report and recommendations*. Canberra: Department of Education, Science and Training.
- Ehri, L. C., Nunes, S. R., Stahl, S. S., & Willows, D. W. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, 71, 393–447.
- Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-to-one tutoring programs for reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, 92, 605–619.
- Esposito, K. (2004). *Reading Recovery program deemed a failure*. Retrieved May 2, 2005, from http://www.nrrf.org/rrp_failure_11-4-04.htm
- Fawcett, A. J., Nicolson, R. I., Moss, H., Nicolson, M. K., & Reason, R. (2001). Effectiveness of reading intervention in junior school. *Educational Psychology*, 21, 299–313.
- Gilmore, A., Croft, C., & Reid, N. (1981). *Burt Word Reading Test—New Zealand revision*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Gomez-Bellenge, F. X., Rodgers, E. M., & Schulz, M. (2005). *Reading Recovery and Descubriendo la Lectura national report 2003–2004*. Columbus, OH: Ohio State University. Retrieved May 26,

- 2005, from <http://www.ndec.us/WebDocs/Documentation/Reading%20Recovery%20and%20Descubriendo%20la%20Lectura%20National%20Report,%202003-2004.pdf>
- Gredler, G. R. (2001). School entrants with learning problems: Assessment and intervention—the American experience. *The Psychology of Education Review*, 25, 22–29.
- Groff, P. (n.d.). *Questions and conclusions from a discussion of Reading Recovery*. Retrieved April 7, 2005, from http://www.nrrf.org/050_disc_of_rr.htm
- Grossen, B., Coulter, G., & Ruggles, B. (1997). *Reading Recovery: An evaluation of benefits and costs* [electronic version]. Retrieved March 21, 2005, from <http://darkwing.uoregon.edu/~bgrossen/rr.htm>
- Haenn, J. F. (2000, April). *Reading Recovery: Success for how many?* Paper presented at the annual meeting of the American Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED442095)
- Hempenstall, K. (1999, February 17). Reading between the lines. *The Age: Education*, p. 5. Retrieved April 25, 2005, from <http://www.rmit.edu.au/departments/ps/staffpgs/hempenart/tween.htm>
- Hiebert, E. (1994). Reading Recovery in the United States: What difference does it make to an age cohort? *Educational Researcher*, 23, 15–25.
- Hobsbaum, A. (1997). Reading Recovery in England. In S. L. Swartz & A. F. Klein (Eds.), *Research in Reading Recovery* (pp. 132–147). Portsmouth, NH: Heinemann.
- Iverson, S., & Tunmer, W. E. (1993). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology*, 85, 112–125.
- Iverson, S., Tunmer, W. E., & Chapman, J. W. (2005). The effects of varying group size on the Reading Recovery approach to preventative early intervention. *Journal of Learning Disabilities*, 38, 456–472.
- Johnston, P., & Allington, R. (1991). Remediation. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 984–1012). White Plains, NY: Longman.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437–447.
- Kerslake, J. (1999). Trends in Reading Recovery data between 1984 and 1998. *The Research Bulletin*, 10, 73–78.
- Kerslake, J. (2000). Annual monitoring of Reading Recovery: The data for 1999. *The Research Bulletin*, 11, 21–26.
- Kerslake, J. (2001). Annual monitoring of Reading Recovery: The data for 2000. *The Research Bulletin*, 12, 65–71.
- Landis, D. (1997). *What about Reading Recovery and effective reading instruction?* Retrieved April 7, 2005, from <http://www.uni.edu/co/ci/file/rr2.html>
- Lyons, C. A. (1997). Reading Recovery and learning disability: Issues, challenges and implications. In S. L. Swartz & A. F. Klein (Eds.), *Research in Reading Recovery* (pp. 122–131). Portsmouth, NH: Heinemann.
- Lyons, C. A. (2003). Reading Recovery in the United States: More than a decade of data. In S. Forbes & C. Briggs (Eds.), *Research in Reading Recovery* (Vol. 2, pp. 215–230). Portsmouth, NH: Heinemann.
- Lyons, C. A., & Beaver, J. (1995). Reducing retention and learning disability placement through Reading Recovery: An educationally sound, cost-effective choice. In R. L. Allington & S. A. Walmsley (Eds.), *No quick fix: Rethinking literacy programs in America's elementary schools* (pp. 116–136). Newark, DE: International Reading Association.
- Ministry of Education. (1999). *Report on the Literacy Taskforce: A report prepared for the Minister of Education*. Wellington, New Zealand: Author.
- Moats, L. C. (2000). *Whole language lives on: The illusion of “balanced reading” instruction*. Retrieved April 15, 2005, from <http://www.edexcellence.net/foundation/publication/publication.cfm?id=45>

- Morris, D., Tyner, B., & Perney, J. (2000). Early Steps: Replicating the effects of a first-grade reading intervention program. *Journal of Educational Psychology*, 92, 681–693.
- Munn, P., & Ellis, S. (2005). Interactions between school systems and Reading Recovery Programmes—evidence from Northern Ireland. *The Curriculum Journal*, 16, 341–362.
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- New South Wales Department of Education and Training. (2003). *2003 Annual Report* [electronic version]. Retrieved April 15, 2005, from https://www.det.nsw.edu.au/media/downloads/reports_stats/annual_reports/year03
- New South Wales Department of Education and Training. (2004). *2004 Annual Report* [electronic version]. Retrieved September 1, 2005, from https://www.det.nsw.edu.au/media/downloads/reports_stats/annual_reports/year04
- Nicolson, R. I., Fawcett, A. J., Moss, H., Nicolson, M. K., & Reason, R. (1999). Early reading intervention can be effective and cost-effective. *British Journal of Educational Psychology*, 69, 47–62.
- Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly*, 29, 9–38.
- Pinnell, G. S., Lyons, C., & Jones, N. (1996). Response to Hiebert: What difference does Reading Recovery make? *Educational Researcher*, 25, 26–28.
- Plewis, I. (2000). Evaluating educational interventions using multilevel growth curves: The case of Reading Recovery. *Educational Research and Evaluation*, 6, 83–101.
- Quay, L. C., Steele, D. C., Johnson, C. I., & Hortman, W. (2003). Children's achievement and personal and social development in a first-year Reading Recovery program with teachers in training. In S. Forbes & C. Briggs (Eds.), *Research in Reading Recovery* (Vol. 2, pp. 281–295). Portsmouth, NH: Heinemann.
- Reading Recovery Council of North America. (2004a). *Reading Recovery celebrates 20 years of accomplishments in North America*. Retrieved May 26, 2005, from <http://www.readingrecovery.org/sections/home/anniversary/highlights.asp>
- Reading Recovery Council of North America. (2004b). *What is implementation?* Retrieved August 31, 2005 from <http://rrcna.com/sections/implementation/index.asp>
- Reading Recovery Council of North America. (2006). *Evidence ignored, learning denied: The attack on Reading Recovery*. Submission to the Inspector General's Office of the U.S. Department of Education [electronic version]. Retrieved 5 April, 2006, from <http://rrcna.com/pdfs/IGevidenceFINAL3-16revWEBATT.pdf>
- Reutzel, D. R. (1999). On Welna's sacred cows: Where's the beef? *The Reading Teacher*, 53, 96–99.
- Reynolds, D. K. (1993). The Reading Recovery program as it relates to understanding. *Reading Improvement*, 30, 76–81.
- Rose, J. (2006). *Independent review of the teaching of early reading: Final report*. London: Department for Education and Skills.
- Rowe, K. J. (1997). Factors affecting students' progress in reading: Key findings from a longitudinal study. In S. L. Swartz & A. F. Klein (Eds.), *Research in Reading Recovery* (pp. 53–101). Portsmouth, NH: Heinemann.
- Schwartz, R. M. (2005). Literacy learning of at-risk first-grade students in the Reading Recovery early intervention. *Journal of Educational Psychology*, 97, 257–267.
- Seethaler, P. M., & Fuchs, L. S. (2005). A drop in the bucket: Randomized controlled trials testing reading and maths interventions. *Learning Disabilities Research & Practice*, 20, 98–102.
- Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly*, 30, 958–996.

- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: The Guilford Press.
- Swartz, S. L., & Klein, A. F. (Eds.). (1997). *Research in Reading Recovery*. Portsmouth, NH: Heinemann.
- Tunmer, W., & Chapman, J. W. (2003). The Reading Recovery approach to preventative early intervention: As good as it gets? *Reading Psychology*, 24, 337–360.
- Victorian Department of Education and Training. (2003). *Reading Recovery: Intervention results*. Retrieved May 29, 2005, from <http://www.sofweb.vic.edu.au/eyes/rr/data.htm>
- Wasik, B. A., & Slavin, R. E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly*, 28, 178–200.
- Welna, L. D. (1999). Balance and sacred cows: A reply to Reutzel. *Reading Teacher*, 53, 94–95.
- Wheldall, K., Center, Y., & Freeman, L. (1992). *Reading Recovery one year on: A follow up evaluation in NSW schools*. Unpublished paper, Macquarie University Special Education Centre, Sydney, Australia.
- Wheldall, K., Center, Y., & Freeman, L. (1993). Reading Recovery in Sydney primary schools. *Australasian Journal of Special Education*, 17, 51–63.

From: Stephanie Metzger (b)(6)
Sent: Monday, June 21, 2010 3:56 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)
WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading

Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto: (b) (6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turnner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporated systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On Mon, 5/24/10, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: (b) (6)
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b) (6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: Stephanie Metzger (b)(6)

Sent: Monday, June 21, 2010 4:08 PM

To: WhatWorks

Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.

<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, Stephanie Metzger (b)(6) wrote:

From: Stephanie Metzger (b)(6)

Subject: RE: What Works Clearinghouse (WWCPC 2177)

To: "WhatWorks" <WWorks@icfi.com>

Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who

knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto: [REDACTED]]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turnner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On Mon, 5/24/10, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: [REDACTED]
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public,

authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies

(<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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:

(b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From:

(b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: WhatWorks
Sent: Tuesday, June 22, 2010 11:04 AM
To: 'Stephanie Metzger'
Subject: RE: What Works Clearinghouse (WWCPC 2230)
Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [REDACTED]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.
<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, **Stephanie Metzger** <steffmetzger@yahoo.com> wrote:

From: Stephanie Metzger [REDACTED]
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually

meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]

Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10**, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: [REDACTED]
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance

with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: Stephanie Metzger [REDACTED]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Follow Up Flag: Follow up

Flag Status: Red

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published it's conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus, Ohio.

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.

2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetic and general reading achievement domains, given results of the Observation Survey. Here are my questions:

a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetic domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetic" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetic domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetic domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all

other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" [REDACTED]
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.

<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On Mon, 6/21/10, Stephanie Metzger (b)(6) wrote:

From: Stephanie Metzger [mailto:stmetzger@icfi.com]
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until

after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto: (b) (6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Tumner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On Mon, 5/24/10, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: (b)(6)
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/idocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: WhatWorks
Sent: Thursday, June 24, 2010 3:50 PM
To: 'Stephanie Metzger'
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:stephanie.metzger@ed.gov]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published its conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus, Ohio.

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.
2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetics and general reading achievement domains, given results of the Observation Survey. Here are my questions:

a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetics domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetics" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetics domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetics domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" <(b) (6)>
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.

<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, Stephanie Metzger <steffmetzger@yahoo.com> wrote:

From: Stephanie Metzger <(b)(6)>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but

rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the WWC Procedures and Standards Handbook, inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto: (b) (6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turnner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10**, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: (b) (6)
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b) (6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:
1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact
link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: Stephanie Metzger (b)(6)
Sent: Tuesday, July 27, 2010 10:56 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2230)

WWC:

I thought I would pass along to you, an article appearing on the RRCNA (Reading Recovery Council of North America) website. It can be easily accessed online simply by going to the RRCNA homepage and clicking on the link about the National Center on Response to Intervention and how it has given Reading Recovery "high marks." I believe that you also ought to know that the RRCNA does seem to misrepresent the conclusions of the WWC. Flyers, articles, and other media publications from the RRCNA appear to exaggerate your actual conclusions. I am reviewing the 1994 Pinnell study at this moment. I do want you to know that I appreciate your replies, and that I intend to continue to communicate with you as I read about the manner in which your ratings were reached. Our previous correspondence appears below. In the meantime, you may want to investigate the RRCNA website. Smoke and mirrors abound.

Stephanie

From: WhatWorks <WWWorks@icfi.com>
To: Stephanie Metzger (b)(6)
Sent: Thu, June 24, 2010 4:50:07 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger (b)(6)
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published its conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus, Ohio .

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.

2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetics and general reading achievement domains, given results of the Observation Survey. Here are my questions:

a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetics domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetics" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetics domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetics domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and

Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.
<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, Stephanie Metzger (b)(6) wrote:

From: Stephanie Metzger <(b)(6)>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of

"outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" <steffmetzger@yahoo.com>
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be

incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto: (b)(6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10, WhatWorks <WhatWorks@icfi.com>** wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: (b)(6)
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public,

authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies

(<http://ies.ed.gov/ncee/wwc/references/ldocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact

link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: Stephanie Metzger (b)(6)
Sent: Tuesday, July 27, 2010 10:57 AM
To: WhatWorks
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

Attachments: NCRTI_Release-7-16-10.pdf

I apologize profusely. I forgot to attach! Here is the article.

----- Forwarded Message -----

From: Stephanie Metzger (b)(6)
To: WhatWorks <WWorks@icfi.com>
Sent: Tue, July 27, 2010 11:55:46 AM
Subject: Re: What Works Clearinghouse (WWCPC 2230)

WWC:

I thought I would pass along to you, an article appearing on the RRCNA (Reading Recovery Council of North America) website. It can be easily accessed online simply by going to the RRCNA homepage and clicking on the link about the National Center on Response to Intervention and how it has given Reading Recovery "high marks." I believe that you also ought to know that the RRCNA does seem to misrepresent the conclusions of the WWC. Flyers, articles, and other media publications from the RRCNA appear to exaggerate your actual conclusions. I am reviewing the 1994 Pinnell study at this moment. I do want you to know that I appreciate your replies, and that I intend to continue to communicate with you as I read about the manner in which your ratings were reached. Our previous correspondence appears below. In the meantime, you may want to investigate the RRCNA website. Smoke and mirrors abound.

Stephanie

From: WhatWorks <WWorks@icfi.com>
To: Stephanie Metzger (b)(6)
Sent: Thu, June 24, 2010 4:50:07 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:stmetzger@ies.ed.gov]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published its conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus , Ohio .

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.

2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetic and general reading achievement domains, given results of the Observation Survey. Here are my questions:

a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetic domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetic" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetic domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetic domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable).

The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" <(b)(6)>

Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.
<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, **Stephanie Metzger** <steffmetzger@yahoo.com> wrote:

From: Stephanie Metzger <(b)(6)>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-

91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>

Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" <(b)(6)>
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On Mon, 5/24/10, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: (b)(6)
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies

(<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact

link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: WhatWorks
Sent: Tuesday, July 27, 2010 12:15 PM
To: 'Stephanie Metzger'
Subject: RE: What Works Clearinghouse (WWCPC 2230)
[Dear Ms. Metzger,](#)

[Thank you for your email and attachment.](#)

[What Works Clearinghouse](#)

From: Stephanie Metzger [mailto:[\(b\)\(6\)](#)]
Sent: Tuesday, July 27, 2010 10:56 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2230)

WWC:

I thought I would pass along to you, an article appearing on the RRCNA (Reading Recovery Council of North America) website. It can be easily accessed online simply by going to the RRCNA homepage and clicking on the link about the National Center on Response to Intervention and how it has given Reading Recovery "high marks." I believe that you also ought to know that the RRCNA does seem to misrepresent the conclusions of the WWC. Flyers, articles, and other media publications from the RRCNA appear to exaggerate your actual conclusions. I am reviewing the 1994 Pinnell study at this moment. I do want you to know that I appreciate your replies, and that I intend to continue to communicate with you as I read about the manner in which your ratings were reached. Our previous correspondence appears below. In the meantime, you may want to investigate the RRCNA website. Smoke and mirrors abound.

Stephanie

From: WhatWorks <WWorks@icfi.com>
To: Stephanie Metzger [\(b\)\(6\)](#)
Sent: Thu, June 24, 2010 4:50:07 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,
Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:stephanie.metzger@ed.gov]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published its conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus , Ohio .

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.
2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetics and general reading achievement domains, given results of the Observation Survey. Here are my questions:

a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetics domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetics" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetics domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetics domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the

majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" <(b)(6)>
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.
<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, **Stephanie Metzger** <steffmetzger@yahoo.com> wrote:

From: Stephanie Metzger <(b)(6)>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I

have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10**, **WhatWorks <WhatWorks@icfi.com>** wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: [REDACTED]
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/idocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:
1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact
link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: Stephanie Metzger (b)(6)
Sent: Tuesday, July 27, 2010 11:54 AM
To: WhatWorks
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

Attachments: NCRTI_Release-7-16-10.pdf

What Works:

I have not been able to peruse thhe WWC's conclusions regarding Reading Recovery for quite some time, but as I have stated below, have not forgotten. This morning, I was able to read about the 1994 Pinnell, Lyons, DeFord, Bryk and Seltzer study. As you might suspect, I have serious concerns and questions. Immediately below I will try to convey them.

1. Of course, Pinnell and her fellow "researchers" are proponents of the Reading Recovery program and its methods. I do understand that the WWC does concern itself with "who conducted" the "research," as long as there are elements of experimental or quasi-experimental methods involved. I find this a serious problem, although I do understand that the number of studies meeting "the gold standard" are rare.
2. I find it interesting that the original "n" was 403 in 43 different schools, yet the actual number of students the WWC considered as it reviewed the study was a mere 79 students. Given that 31 students were randomly assigned to the "treatment" or intervention group and 48 were randomly assigned to "no intervention," and the actual study reported an original "n" of 403, it appears, on its face, that the WWC found it completely necessary to "dig" through the study to attempt to locate any element of random assignment at all. It is almost amusing that the authors of the study didn't find it necessary to compare students randomly assigned to Reading Recovery to those who received other types of interventions (Direct Instruction; Reading/Writing Group; Reading Successfully). True researchers understand that random assignment occurs across different treatment groups, and that effect sizes are calculated for each. I apologize, but the study should be thrown out on its face.
3. So, here were are. We have found at least some elements of the "study" that appeared sound, with a total "n" of 79 (31 who got the intervention and 48 who did not). It is now almost laughable that the 48 who did not receive the treatment, were children who received the "regular reading program." And what, pray tell, would that be? The children in the comparison group "did not receive any special instruction." Most educators know that anything is better than nothing at all. Time on task alone had to be a factor. Was time on task controlled? Have the WWC

reviewers heard of the Matthew Effect cited in the literature? How are we to ever know whether the children in the control group were even exposed to daily, small group instruction, or that the time they spent engaged in "real reading" was comparable to the 30-minutes-a-day requirement of Reading Recovery? Was the "regular program" described in the study?

4. Once again, the dictation subtest (Hearing and Recording Sounds in Words) yielded an effect size of .65, but we all know that this subtest does, and never will, measure "general reading achievement." It's incredible to me that the WWC would allow this effect size to even be considered in a rating along this construct. Also, why did the WWC not calculate a rating for the comprehension construct if the Gates MacGinitie Reading Test truly measures vocabulary and comprehension as described? Similarly, since I am familiar with the Woodcock Reading Mastery Test-R, I am wondering why there was no calculation of effect size for alphabets? Did the WWC simply decide to toss everything and just look at general reading achievement based on results across three assessments that all measure totally different things? This is unbelievable.

5. How did the study control for training that teachers had? The review alone demonstrates that the Reading Recovery teachers all had a certain amount of "training." Did the authors describe what kind of "training" the comparison group teachers had? Were they entry-level teachers? Hopefully, they were not paraprofessionals who delivered a regular reading program (no intervention) that I asked about in a previous email.

Thank you for your time.

Stephanie Metzger

----- Forwarded Message -----

From: Stephanie Metzger (b)(6)
To: WWorks@icfi.com
Sent: Tue, July 27, 2010 11:57:25 AM
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

I apologize profusely. I forgot to attach! Here is the article.

----- Forwarded Message -----

From: Stephanie Metzger (b)(6)
To: WhatWorks <WWorks@icfi.com>
Sent: Tue, July 27, 2010 11:55:46 AM
Subject: Re: What Works Clearinghouse (WWCPC 2230)

WWC:

I thought I would pass along to you, an article appearing on the RRCNA (Reading Recovery Council of North America) website. It can be easily accessed online simply by going to the RRCNA homepage and clicking on the link about the National Center on Response to Intervention and how it has given Reading Recovery "high marks." I believe that you also ought to know that the RRCNA does seem to misrepresent the conclusions of the WWC. Flyers, articles, and other media publications from the RRCNA appear to exaggerate your actual conclusions. I am reviewing the 1994 Pinnell study at this moment. I do want you to know that I appreciate your replies, and that I intend to continue to communicate with you as I read about the manner in which your ratings were reached. Our previous correspondence appears below. In the meantime, you may want to investigate the RRCNA website. Smoke and mirrors abound.

Stephanie

From: WhatWorks <WWorks@icfi.com>
To: Stephanie Metzger [mailto:stephanie@icfi.com]
Sent: Thu, June 24, 2010 4:50:07 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:stephanie@icfi.com]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published it's conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus , Ohio .

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.

2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetic and general reading achievement domains, given results of the Observation Survey. Here are my questions:

a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetic domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetic" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetic domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetic domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the

student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" [REDACTED]
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.
<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, **Stephanie Metzger** <steffmetzger@yahoo.com> wrote:

From: Stephanie Metzger <(b)(6)>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporated systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10**, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: (b)(6)
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies

(<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact

link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: WhatWorks

Sent: Tuesday, July 27, 2010 12:18 PM

To: 'Stephanie Metzger'

Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions. Please let us know when you have submitted all of your questions so that the QRT team can begin to work on a response.

Thank you,

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/>.

From: Stephanie Metzger [mailto: (b) (6)]
Sent: Tuesday, July 27, 2010 11:54 AM
To: WhatWorks
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

What Works:

I have not been able to peruse thhe WWC's conclusions regarding Reading Recovery for quite some time, but as I have stated below, have not forgotten. This morning, I was able to read about the 1994 Pinnell, Lyons, DeFord, Bryk and Seltzer study. As you might suspect, I have serious concerns and questions. Immediately below I will try to convey them.

1. Of course, Pinnell and her fellow "researchers" are proponents of the Reading Recovery program and its methods. I do understand that the WWC does concern itself with "who conducted" the "research," as long as there are elements of experimental or quasi-experimental methods involved. I find this a serious problem, although I do understand that the number of studies meeting "the gold standard" are rare.

2. I find it interesting that the original "n" was 403 in 43 different schools, yet the actual number of students the WWC considered as it reviewed the study was a mere 79 students. Given that 31 students were randomly

assigned to the "treatment" or intervention group and 48 were randomly assigned to "no intervention," and the actual study reported an original "n" of 403, it appears, on its face, that the WWC found it completely necessary to "dig" through the study to attempt to locate any element of random assignment at all. It is almost amusing that the authors of the study didn't find it necessary to compare students randomly assigned to Reading Recovery to those who received other types of interventions (Direct Instruction; Reading/Writing Group; Reading Successfully). True researchers understand that random assignment occurs across different treatment groups, and that effect sizes are calculated for each. I apologize, but the study should be thrown out on its face.

3. So, here we are. We have found at least some elements of the "study" that appeared sound, with a total "n" of 79 (31 who got the intervention and 48 who did not). It is now almost laughable that the 48 who did not receive the treatment, were children who received the "regular reading program." And what, pray tell, would that be? The children in the comparison group "did not receive any special instruction." Most educators know that anything is better than nothing at all. Time on task alone had to be a factor. Was time on task controlled? Have the WWC reviewers heard of the Matthew Effect cited in the literature? How are we to ever know whether the children in the control group were even exposed to daily, small group instruction, or that the time they spent engaged in "real reading" was comparable to the 30-minutes-a-day requirement of Reading Recovery? Was the "regular program" described in the study?

4. Once again, the dictation subtest (Hearing and Recording Sounds in Words) yielded an effect size of .65, but we all know that this subtest does, and never will, measure "general reading achievement." It's incredible to me that the WWC would allow this effect size to even be considered in a rating along this construct. Also, why did the WWC not calculate a rating for the comprehension construct if the Gates MacGinitie Reading Test truly measures vocabulary and comprehension as described? Similarly, since I am familiar with the Woodcock Reading Mastery Test-R, I am wondering why there was no calculation of effect size for alphabets? Did the WWC simply decide to toss everything and just look at general reading achievement based on results across three assessments that all measure totally different things? This is unbelievable.

5. How did the study control for training that teachers had? The review alone demonstrates that the Reading Recovery teachers all had a certain amount of "training." Did the authors describe what kind of "training" the comparison group teachers had? Were they entry-level teachers? Hopefully, they were not paraprofessionals who delivered a regular reading program (no intervention) that I asked about in a previous email.

Thank you for your time.

Stephanie Metzger

----- Forwarded Message -----

From: Stephanie Metzger (b)(6)
To: WWorks@icfi.com
Sent: Tue, July 27, 2010 11:57:25 AM
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

I apologize profusely. I forgot to attach! Here is the article.

----- Forwarded Message -----

From: Stephanie Metzger (b)(6)
To: WhatWorks <WWorks@icfi.com>
Sent: Tue, July 27, 2010 11:55:46 AM
Subject: Re: What Works Clearinghouse (WWCPC 2230)

WWC:

I thought I would pass along to you, an article appearing on the RRCNA (Reading Recovery Council of North America) website. It can be easily accessed online simply by going to the RRCNA homepage and clicking on the link about the National Center on Response to Intervention and how it has given Reading Recovery "high marks." I believe that you also ought to know that the RRCNA does seem to misrepresent the conclusions of the WWC. Flyers, articles, and other media publications from the RRCNA appear to exaggerate your actual conclusions. I am reviewing the 1994 Pinnell study at this moment. I do want you to know that I appreciate your replies, and that I intend to continue to communicate with you as I read about the manner in which your ratings were reached. Our previous correspondence appears below. In the meantime, you may want to investigate the RRCNA website. Smoke and mirrors abound.

Stephanie

From: WhatWorks <WWorks@icfi.com>
To: Stephanie Metzger (b)(6)
Sent: Thu, June 24, 2010 4:50:07 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published it's conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus , Ohio .

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.
2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetics and general reading achievement domains, given results of the Observation Survey. Here are my questions:

- a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetics domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that

require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetics" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetics domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetics domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate

measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.
<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, **Stephanie Metzger** <steffmetzger@yahoo.com> wrote:

From: Stephanie Metzger <(b)(6)>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small

groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" <(b)(6)>
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]

Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10**, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: (b)(6)
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence

standards used when reviewing studies

(<http://ies.ed.gov/ncee/wwc/references/ldocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:
1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact
link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: Stephanie Metzger (b)(6)
Sent: Thursday, July 29, 2010 10:25 AM
To: WhatWorks
Subject: Fw: What Works Clearinghouse (WWCPC 2230)
What Works:

I appreciate all of your prompt replies to my extended series of emails. My questions/concerns below pertain to the Schwartz study conducted in 2005.

1. R.M. Schwartz authored, it appears, an article about "the history of Reading Recovery" and the "legacy of Marie Clay." Although the WWC did not consider this to be a study (and rightly so), it seems, right off the bat, that R.M. Schwartz is a "fan" of Reading Recovery. Again, as in previous emails, I raise concern that there is inherent bias on the part of the "researchers" who submitted their studies for consideration (i.e. Pinnell and company, Schwartz).

2. Students for the treatment (n=37) and comparison (n=37) were, as indicated, identified as "eligible" for Reading Recovery on the basis of their performance on 6 subtests of the Observation Survey and "teacher judgment." Teacher judgment was a factor in determining eligibility? Doesn't sound very objective to me. How were the students "randomly" assigned to the treatment and comparison groups? Was "teacher judgment" involved in the "random" assignment, or was the random assignment done through flip of the coin? How was the "random assignment" controlled?

3. Along the alphabetics domain, phonemic awareness, print awareness, letter identification and phonics construct effect sizes are reported by Schwartz. So, the effect on phonemic awareness was not statistically significant, based on the 10-item deletion task and the Yopp-Singer results. This is not surprising. Similarly, no statistically significant effect was found for letter identification (as per the WWC in contrast to the conclusion on the part of the researcher). The WWC confirmed that on two constructs (print awareness and "phonics") effect sizes were statistically significant, yet "print awareness" results were measured through the "concepts about print" subtest (my concerns with this subtest are expressed in a previous email) and "phonics" or phonetic skills were measured by the word recognition subtest from the Observation Survey. Again, the word recognition subtest does not measure whether a child has solid knowledge/awareness of the six syllable types (patterns) of English. One cannot rule out the fact that some words (e.g. the word "there") are merely "sight words." For the WWC to conclude that a word recognition subtest is a true measure of "phonics" tells me that this

particular fact was either ignored, or that it was not understood by the WWC content experts and principal investigator.

4. It is not surprising, once again, that the children in the 1st round of Reading Recovery (treatment group) seemed to outperform the children who had to wait until 2nd round. While it appears, on its face, that the skills of the two groups were similar (matched) prior to the treatment, it is (again) clear that the children on the "wait list" simply "waited," getting only the regular classroom reading instruction. In other words, no control for "time on task" was indicated in the WWC's review of the study. Again, any treatment that leads to more time engaged in learning is bound to yield positive effects. Could it be that the "time on task" factor alone, resulted in seemingly better results for the children in the treatment group?

5. As I stated in a previous email, the dictation and writing vocabulary subtests from the Observation Survey are not measures of general reading achievement!! They require auditory memory (dictation subtest) and strong spelling skills (both subtests).

Thank you for your time. Also know that I will be sending one more email with questions.

Stephanie Metzger

----- Forwarded Message -----

From: WhatWorks <WWorks@icfi.com>

To: Stephanie Metzger (b)(6)

Sent: Tue, July 27, 2010 1:18:03 PM

Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions. Please let us know when you have submitted all of your questions so that the QRT team can begin to work on a response.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Tuesday, July 27, 2010 11:54 AM
To: WhatWorks
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

What Works:

I have not been able to peruse thhe WWC's conclusions regarding Reading Recovery for quite some time, but as I have stated below, have not forgotten. This morning, I was able to read about the 1994 Pinnell, Lyons , DeFord, Bryk and Seltzer study. As you might suspect, I have serious concerns and questions. Immediately below I will try to convey them.

1. Of course, Pinnell and her fellow "researchers" are proponents of the Reading Recovery program and its methods. I do understand that the WWC does concern itself with "who conducted" the "research," as long as there are elements of experimental or quasi-experimental methods involved. I find this a serious problem, although I do understand that the number of studies meeting "the gold standard" are rare.
2. I find it interesting that the original "n" was 403 in 43 different schools, yet the actual number of students the WWC considered as it reviewed the study was a mere 79 students. Given that 31 students were randomly assigned to the "treatment" or intervention group and 48 were randomly assigned to "no intervention," and the actual study reported an original "n" of 403, it appears, on its face, that the WWC found it completely necessary to "dig" through the study to attempt to locate any element of random assignment at all. It is almost amusing that the authors of the study didn't find it necessary to compare students randomly assigned to Reading Recovery to those who received other types of interventions (Direct Instruction; Reading/Writing Group; Reading Successfully). True researchers understand that random assignment occurs across different treatment groups, and that effect sizes are calculated for each. I apologize, but the study should be thrown out on its face.
3. So, here were are. We have found at least some elements of the "study" that appeared sound, with a total "n" of 79 (31 who got the intervention and 48 who did not). It is now almost laughable that the 48 who did not receive the treatment, were children who received the "regular reading program." And what, pray tell, would that be? The children in the comparison group "did not receive any special instruction." Most educators know that anything is better than nothing at all. Time on task alone had to be a factor. Was time on task controlled? Have the WWC

reviewers heard of the Matthew Effect cited in the literature? How are we to ever know whether the children in the control group were even exposed to daily, small group instruction, or that the time they spent engaged in "real reading" was comparable to the 30-minutes-a-day requirement of Reading Recovery? Was the "regular program" described in the study?

4. Once again, the dictation subtest (Hearing and Recording Sounds in Words) yielded an effect size of .65, but we all know that this subtest does, and never will, measure "general reading achievement." It's incredible to me that the WWC would allow this effect size to even be considered in a rating along this construct. Also, why did the WWC not calculate a rating for the comprehension construct if the Gates MacGinitie Reading Test truly measures vocabulary and comprehension as described? Similarly, since I am familiar with the Woodcock Reading Mastery Test-R, I am wondering why there was no calculation of effect size for alphabets? Did the WWC simply decide to toss everything and just look at general reading achievement based on results across three assessments that all measure totally different things? This is unbelievable.

5. How did the study control for training that teachers had? The review alone demonstrates that the Reading Recovery teachers all had a certain amount of "training." Did the authors describe what kind of "training" the comparison group teachers had? Were they entry-level teachers? Hopefully, they were not paraprofessionals who delivered a regular reading program (no intervention) that I asked about in a previous email.

Thank you for your time.

Stephanie Metzger

----- Forwarded Message -----

From: Stephanie Metzger <steffmetzger@yahoo.com>
To: WWorks@icfi.com
Sent: Tue, July 27, 2010 11:57:25 AM
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

I apologize profusely. I forgot to attach! Here is the article.

----- Forwarded Message -----

From: Stephanie Metzger <steffmetzger@yahoo.com>
To: WhatWorks <WWorks@icfi.com>
Sent: Tue, July 27, 2010 11:55:46 AM
Subject: Re: What Works Clearinghouse (WWCPC 2230)

WWC:

I thought I would pass along to you, an article appearing on the RRCNA (Reading Recovery Council of North America) website. It can be easily accessed online simply by going to the RRCNA homepage and clicking on the link about the National Center on Response to Intervention and how it has given Reading Recovery "high marks." I believe that you also ought to know that the RRCNA does seem to misrepresent the conclusions of the WWC. Flyers, articles, and other media publications from the RRCNA appear to exaggerate your actual conclusions. I am reviewing the 1994 Pinnell study at this moment. I do want you to know that I appreciate your replies, and that I intend to continue to communicate with you as I read about the manner in which your ratings were reached. Our previous correspondence appears below. In the meantime, you may want to investigate the RRCNA website. Smoke and mirrors abound.

Stephanie

From: WhatWorks <WWorks@icfi.com>
To: Stephanie Metzger (b)(6)
Sent: Thu, June 24, 2010 4:50:07 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published it's conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will

consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus , Ohio .

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.

2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetic and general reading achievement domains, given results of the Observation Survey. Here are my questions:

a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetic domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetic" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetic domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetic domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading. although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a

measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.
<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, Stephanie Metzger (b)(6) wrote:

From: Stephanie Metzger (b)(6)
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>

Subject: RE: What Works Clearinghouse (WWCPC 2177)

To: "Stephanie Metzger" [REDACTED]

Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On Mon, 5/24/10, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: (b)(6)
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to

check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies

(<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact

link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm


From: WhatWorks
Sent: Thursday, July 29, 2010 10:32 AM
To: 'Stephanie Metzger'
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for your email.

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/>.

From: Stephanie Metzger [mailto:
Sent: Thursday, July 29, 2010 10:25 AM
To: WhatWorks
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

What Works:

I appreciate all of your prompt replies to my extended series of emails. My questions/concerns below pertain to the Schwartz study conducted in 2005.

1. R.M. Schwartz authored, it appears, an article about "the history of Reading Recovery" and the "legacy of Marie Clay." Although the WWC did not consider this to be a study (and rightly so), it seems, right off the bat, that R.M. Schwartz is a "fan" of Reading Recovery. Again, as in previous emails, I raise concern that there is inherent bias on the part of the "researchers" who submitted their studies for consideration (i.e. Pinnell and company, Schwartz).

2. Students for the treatment (n=37) and comparison (n=37) were, as indicated, identified as "eligible" for Reading Recovery on the basis of their performance on 6 subtests of the Observation Survey and "teacher judgment." Teacher judgment was a factor in determining eligibility? Doesn't sound very objective to me. How were the students "randomly" assigned to the treatment and comparison groups? Was "teacher judgment" involved in the "random" assignment, or was the random assignment done through flip of the coin? How was the "random assignment" controlled?

3. Along the alphabetics domain, phonemic awareness, print awareness, letter identification and phonics construct effect sizes are reported by

Schwartz. So, the effect on phonemic awareness was not statistically significant, based on the 10-item deletion task and the Yopp-Singer results. This is not surprising. Similarly, no statistically significant effect was found for letter identification (as per the WWC in contrast to the conclusion on the part of the researcher). The WWC confirmed that on two constructs (print awareness and "phonics") effect sizes were statistically significant, yet "print awareness" results were measured through the "concepts about print" subtest (my concerns with this subtest are expressed in a previous email) and "phonics" or phonetic skills were measured by the word recognition subtest from the Observation Survey. Again, the word recognition subtest does not measure whether a child has solid knowledge/awareness of the six syllable types (patterns) of English. One cannot rule out the fact that some words (e.g. the word "there") are merely "sight words." For the WWC to conclude that a word recognition subtest is a true measure of "phonics" tells me that this particular fact was either ignored, or that it was not understood by the WWC content experts and principal investigator.

4. It is not surprising, once again, that the children in the 1st round of Reading Recovery (treatment group) seemed to outperform the children who had to wait until 2nd round. While it appears, on its face, that the skills of the two groups were similar (matched) prior to the treatment, it is (again) clear that the children on the "wait list" simply "waited," getting only the regular classroom reading instruction. In other words, no control for "time on task" was indicated in the WWC's review of the study. Again, any treatment that leads to more time engaged in learning is bound to yield positive effects. Could it be that the "time on task" factor alone, resulted in seemingly better results for the children in the treatment group?

5. As I stated in a previous email, the dictation and writing vocabulary subtests from the Observation Survey are not measures of general reading achievement!! They require auditory memory (dictation subtest) and strong spelling skills (both subtests).

Thank you for your time. Also know that I will be sending one more email with questions.

Stephanie Metzger

----- Forwarded Message -----

From: WhatWorks <WWorks@icfi.com>

To: Stephanie Metzger

(b)(6)

Sent: Tue, July 27, 2010 1:18:03 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions. Please let us know when you have submitted all of your questions so that the QRT team can begin to work on a response.

Thank you,

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/>.

From: Stephanie Metzger [mailto: (b) (6)]
Sent: Tuesday, July 27, 2010 11:54 AM
To: WhatWorks
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

What Works:

I have not been able to peruse thhe WWC's conclusions regarding Reading Recovery for quite some time, but as I have stated below, have not forgotten. This morning, I was able to read about the 1994 Pinnell, Lyons , DeFord, Bryk and Seltzer study. As you might suspect, I have serious concerns and questions. Immediately below I will try to convey them.

1. Of course, Pinnell and her fellow "researchers" are proponents of the Reading Recovery program and its methods. I do understand that the WWC does concern itself with "who conducted" the "research," as long as there are elements of experimental or quasi-experimental methods involved. I find this a serious problem, although I do understand that the number of studies meeting "the gold standard" are rare.

2. I find it interesting that the original "n" was 403 in 43 different schools, yet the actual number of students the WWC considered as it reviewed the study was a mere 79 students. Given that 31 students were randomly assigned to the "treatment" or intervention group and 48 were randomly

assigned to "no intervention," and the actual study reported an original "n" of 403, it appears, on its face, that the WWC found it completely necessary to "dig" through the study to attempt to locate any element of random assignment at all. It is almost amusing that the authors of the study didn't find it necessary to compare students randomly assigned to Reading Recovery to those who received other types of interventions (Direct Instruction; Reading/Writing Group; Reading Successfully). True researchers understand that random assignment occurs across different treatment groups, and that effect sizes are calculated for each. I apologize, but the study should be thrown out on its face.

3. So, here we are. We have found at least some elements of the "study" that appeared sound, with a total "n" of 79 (31 who got the intervention and 48 who did not). It is now almost laughable that the 48 who did not receive the treatment, were children who received the "regular reading program." And what, pray tell, would that be? The children in the comparison group "did not receive any special instruction." Most educators know that anything is better than nothing at all. Time on task alone had to be a factor. Was time on task controlled? Have the WWC reviewers heard of the Matthew Effect cited in the literature? How are we to ever know whether the children in the control group were even exposed to daily, small group instruction, or that the time they spent engaged in "real reading" was comparable to the 30-minutes-a-day requirement of Reading Recovery? Was the "regular program" described in the study?

4. Once again, the dictation subtest (Hearing and Recording Sounds in Words) yielded an effect size of .65, but we all know that this subtest does, and never will, measure "general reading achievement." It's incredible to me that the WWC would allow this effect size to even be considered in a rating along this construct. Also, why did the WWC not calculate a rating for the comprehension construct if the Gates MacGinitie Reading Test truly measures vocabulary and comprehension as described? Similarly, since I am familiar with the Woodcock Reading Mastery Test-R, I am wondering why there was no calculation of effect size for alphabets? Did the WWC simply decide to toss everything and just look at general reading achievement based on results across three assessments that all measure totally different things? This is unbelievable.

5. How did the study control for training that teachers had? The review alone demonstrates that the Reading Recovery teachers all had a certain amount of

"training." Did the authors describe what kind of "training" the comparison group teachers had? Were they entry-level teachers? Hopefully, they were not paraprofessionals who delivered a regular reading program (no intervention) that I asked about in a previous email.

Thank you for your time.

Stephanie Metzger

----- Forwarded Message -----

From: Stephanie Metzger (b)(6)
To: WWorks@icfi.com
Sent: Tue, July 27, 2010 11:57:25 AM
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

I apologize profusely. I forgot to attach! Here is the article.

----- Forwarded Message -----

From: Stephanie Metzger (b)(6)
To: WhatWorks <WWorks@icfi.com>
Sent: Tue, July 27, 2010 11:55:46 AM
Subject: Re: What Works Clearinghouse (WWCPC 2230)

WWC:

I thought I would pass along to you, an article appearing on the RRCNA (Reading Recovery Council of North America) website. It can be easily accessed online simply by going to the RRCNA homepage and clicking on the link about the National Center on Response to Intervention and how it has given Reading Recovery "high marks." I believe that you also ought to know that the RRCNA does seem to misrepresent the conclusions of the WWC. Flyers, articles, and other media publications from the RRCNA appear to exaggerate your actual conclusions. I am reviewing the 1994 Pinnell study at this moment. I do want you to know that I appreciate your replies, and that I intend to continue to communicate with you as I read about the manner in which your ratings were reached. Our previous correspondence appears below. In the meantime, you may want to investigate the RRCNA website. Smoke and mirrors abound.

Stephanie

From: WhatWorks <WWorks@icfi.com>
To: Stephanie Metzger (b)(6)

Sent: Thu, June 24, 2010 4:50:07 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto: (b)(6)]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published it's conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus , Ohio .

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.
2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetics and general reading achievement domains, given results of the Observation Survey. Here are my questions:

- a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetics domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetics" as defined by the

National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetics domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetics domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first

round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.
<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, **Stephanie Metzger** <steffmetzger@yahoo.com> wrote:

From: Stephanie Metzger <[REDACTED]>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "What Works" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year

assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10, WhatWorks <WhatWorks@icfi.com>** wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: [REDACTED]
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/iddocviewer/doc.aspx?docid=19&tocid=1>). Please

note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: WhatWorks
Sent: Monday, October 18, 2010 3:26 PM
To: 'Stephanie Metzger'
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Our Quality Review Team (QRT) plans to begin preparing a response to your previously submitted questions from June and July 2010 (included below). If you have additional questions that you would like to be considered at this time, please submit them in response to this email.

Thank you,

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/>.

From: WhatWorks
Sent: Thursday, July 29, 2010 10:32 AM
To: 'Stephanie Metzger'
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for your email.

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/>.

From: Stephanie Metzger [mailto: (b) (6)]
Sent: Thursday, July 29, 2010 10:25 AM
To: WhatWorks
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

What Works:

I appreciate all of your prompt replies to my extended series of emails. My questions/concerns below pertain to the Schwartz study conducted in 2005.

1. R.M. Schwartz authored, it appears, an article about "the history of Reading Recovery" and the "legacy of Marie Clay." Although the WWC did not consider this to be a study (and rightly so), it seems, right off the bat, that R.M. Schwartz is a "fan" of Reading Recovery. Again, as in previous emails, I raise concern that there is inherent bias on the part of the "researchers" who submitted their studies for consideration (i.e. Pinnell and company, Schwartz).

2. Students for the treatment (n=37) and comparison (n=37) were, as indicated, identified as "eligible" for Reading Recovery on the basis of their performance on 6 subtests of the Observation Survey and "teacher judgment." Teacher judgment was a factor in determining eligibility? Doesn't sound very objective to me. How were the students "randomly" assigned to the treatment and comparison groups? Was "teacher judgment" involved in the "random" assignment, or was the random assignment done through flip of the coin? How was the "random assignment" controlled?

3. Along the alphabets domain, phonemic awareness, print awareness, letter identification and phonics construct effect sizes are reported by Schwartz. So, the effect on phonemic awareness was not statistically significant, based on the 10-item deletion task and the Yopp-Singer results. This is not surprising. Similarly, no statistically significant effect was found for letter identification (as per the WWC in contrast to the conclusion on the part of the researcher). The WWC confirmed that on two constructs (print awareness and "phonics") effect sizes were statistically significant, yet "print awareness" results were measured through the "concepts about print" subtest (my concerns with this subtest are expressed in a previous email) and "phonics" or phonetic skills were measured by the word recognition subtest from the Observation Survey. Again, the word recognition subtest does not measure whether a child has solid knowledge/awareness of the six syllable types (patterns) of English. One cannot rule out the fact that some words (e.g. the word "there") are merely "sight words." For the WWC to conclude that a word recognition subtest is a true measure of "phonics" tells me that this particular fact was either ignored, or that it was not understood by the WWC content experts and principal investigator.

4. It is not surprising, once again, that the children in the 1st round of Reading Recovery (treatment group) seemed to outperform the children who had to wait until 2nd round. While it appears, on its face, that the skills of the two groups were similar (matched) prior to the treatment, it is (again)

clear that the children on the "wait list" simply "waited," getting only the regular classroom reading instruction. In other words, no control for "time on task" was indicated in the WWC's review of the study. Again, any treatment that leads to more time engaged in learning is bound to yield positive effects. Could it be that the "time on task" factor alone, resulted in seemingly better results for the children in the treatment group?

5. As I stated in a previous email, the dictation and writing vocabulary subtests from the Observation Survey are not measures of general reading achievement!! They require auditory memory (dictation subtest) and strong spelling skills (both subtests).

Thank you for your time. Also know that I will be sending one more email with questions.

Stephanie Metzger

----- Forwarded Message -----

From: WhatWorks <WWorks@icfi.com>

To: Stephanie Metzger (b)(6)

Sent: Tue, July 27, 2010 1:18:03 PM

Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions. Please let us know when you have submitted all of your questions so that the QRT team can begin to work on a response.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]

Sent: Tuesday, July 27, 2010 11:54 AM

To: WhatWorks

Subject: Fw: What Works Clearinghouse (WWCPC 2230)

What Works:

I have not been able to peruse thhe WWC's conclusions regarding Reading Recovery for quite some time, but as I have stated below, have not forgotten. This morning, I was able to read about the 1994 Pinnell, Lyons , DeFord, Bryk and Seltzer study. As you might suspect, I have serious concerns and questions. Immediately below I will try to convey them.

1. Of course, Pinnell and her fellow "researchers" are proponents of the Reading Recovery program and its methods. I do understand that the WWC does concern itself with "who conducted" the "research," as long as there are elements of experimental or quasi-experimental methods involved. I find this a serious problem, although I do understand that the number of studies meeting "the gold standard" are rare.

2. I find it interesting that the original "n" was 403 in 43 different schools, yet the actual number of students the WWC considered as it reviewed the study was a mere 79 students. Given that 31 students were randomly assigned to the "treatment" or intervention group and 48 were randomly assigned to "no intervention," and the actual study reported an original "n" of 403, it appears, on its face, that the WWC found it completely necessary to "dig" through the study to attempt to locate any element of random assignment at all. It is almost amusing that the authors of the study didn't find it necessary to compare students randomly assigned to Reading Recovery to those who received other types of interventions (Direct Instruction; Reading/Writing Group; Reading Successfully). True researchers understand that random assignment occurs across different treatment groups, and that effect sizes are calculated for each. I apologize, but the study should be thrown out on its face.

3. So, here were are. We have found at least some elements of the "study" that appeared sound, with a total "n" of 79 (31 who got the intervention and 48 who did not). It is now almost laughable that the 48 who did not receive the treatment, were children who received the "regular reading program." And what, pray tell, would that be? The children in the comparison group "did not receive any special instruction." Most educators know that anything is better than nothing at all. Time on task alone had to be a factor. Was time on task controlled? Have the WWC reviewers heard of the Matthew Effect cited in the literature? How are we to ever know whether the children in the control group were even exposed to daily, small group instruction, or that

the time they spent engaged in "real reading" was comparable to the 30-minutes-a-day requirement of Reading Recovery? Was the "regular program" described in the study?

4. Once again, the dictation subtest (Hearing and Recording Sounds in Words) yielded an effect size of .65, but we all know that this subtest does, and never will, measure "general reading achievement." It's incredible to me that the WWC would allow this effect size to even be considered in a rating along this construct. Also, why did the WWC not calculate a rating for the comprehension construct if the Gates MacGinitie Reading Test truly measures vocabulary and comprehension as described? Similarly, since I am familiar with the Woodcock Reading Mastery Test-R, I am wondering why there was no calculation of effect size for alphabets? Did the WWC simply decide to toss everything and just look at general reading achievement based on results across three assessments that all measure totally different things? This is unbelievable.

5. How did the study control for training that teachers had? The review alone demonstrates that the Reading Recovery teachers all had a certain amount of "training." Did the authors describe what kind of "training" the comparison group teachers had? Were they entry-level teachers? Hopefully, they were not paraprofessionals who delivered a regular reading program (no intervention) that I asked about in a previous email.

Thank you for your time.

Stephanie Metzger

----- Forwarded Message -----

From: Stephanie Metzger <steffmetzger@yahoo.com>
To: WWorks@icfi.com
Sent: Tue, July 27, 2010 11:57:25 AM
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

I apologize profusely. I forgot to attach! Here is the article.

----- Forwarded Message -----

From: Stephanie Metzger <steffmetzger@yahoo.com>
To: WhatWorks <WWorks@icfi.com>
Sent: Tue, July 27, 2010 11:55:46 AM
Subject: Re: What Works Clearinghouse (WWCPC 2230)

WWC:

I thought I would pass along to you, an article appearing on the RRCNA (Reading Recovery Council of North America) website. It can be easily accessed online simply by going to the RRCNA homepage and clicking on the link about the National Center on Response to Intervention and how it has given Reading Recovery "high marks." I believe that you also ought to know that the RRCNA does seem to misrepresent the conclusions of the WWC. Flyers, articles, and other media publications from the RRCNA appear to exaggerate your actual conclusions. I am reviewing the 1994 Pinnell study at this moment. I do want you to know that I appreciate your replies, and that I intend to continue to communicate with you as I read about the manner in which your ratings were reached. Our previous correspondence appears below. In the meantime, you may want to investigate the RRCNA website. Smoke and mirrors abound.

Stephanie

From: WhatWorks <WWWorks@icfi.com>
To: Stephanie Metzger [mailto: (b)(6)]
Sent: Thu, June 24, 2010 4:50:07 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto: (b)(6)]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published it's conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus , Ohio .

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.

2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetics and general reading achievement domains, given results of the Observation Survey. Here are my questions:

a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetics domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetics" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetics domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetics domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must

remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.
<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10**, **Stephanie Metzger** <steffmetzger@yahoo.com> wrote:

From: Stephanie Metzger <(b)(6)>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks" <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" <[REDACTED]>
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a

Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turnner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and it's proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On Mon, 5/24/10, WhatWorks <WhatWorks@icfi.com> wrote:

From: WhatWorks <WhatWorks@icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: (b)(6)

Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies (<http://ies.ed.gov/ncee/wwc/references/idoctrviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: Stephanie Metzger [s (b)(6)]
Sent: Friday, October 22, 2010 5:42 PM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2230)

Follow Up Flag: Follow up
Flag Status: Red
Dear Quality Review Team:

Just today, I looked at the Iverson and Tunner study that met your standards with reservations. I apologize for not having had the time to look at it sooner. I am extremely busy attempting to "pick up the pieces," designing rapid naming charts and connected text for students who received RR last year (they unfortunately, are not faring well). I also was passed along a copy of the Observation Survey, and found it to be lacking in validity, with regard to its measure of "phonemic awareness," so important in building a foundation for children.

Of course, the study mentioned above had a clear intent. That intent was to determine if adding an instructional piece to the standard Reading Recovery program would result in better, more profound, knowledge on the part of the children studied, in both phonemic awareness and the alphabetic principle. Unfortunately, the children who received the "modified" version of Reading Recovery were not included in the final "analysis."

I do believe that it is important for the WWC quality review team to understand the intent of any study submitted. Again, the "comparison group" received "the regular program," which, to the reader, is not clearly defined. The intent of the study conducted was, really, to compare a standard RR program to that which incorporated a truly important piece of the instructional puzzle...and this, I found to be "tossed out" in the WWC's analysis.

The Observation Survey...especially the Hearing Sounds In Words subtest is an absolute abominable measure....and I would hope that the WWC would consider this as it continues to attempt to understand the nature of "how children learn to read." Did the WWC look at this subtest at all? Louisa Cook Moats would (should) have a field day with this.

If the WWC concludes that the Observation Survey is a valid measure of "the alphabetic principle," "phonemic awareness," or, even better yet, "general reading achievement," then I will take my ball and go home.

Done.

Stephanie Metzger

From: WhatWorks <WhatWorks@icfi.com>
To: Stephanie Metzger (b)(6)
Sent: Mon, October 18, 2010 4:25:31 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Our Quality Review Team (QRT) plans to begin preparing a response to your previously submitted questions from June and July 2010 (included below). If you have additional questions that you would like to be considered at this time, please submit them in response to this email.

Thank you,

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/>.

From: WhatWorks
Sent: Thursday, July 29, 2010 10:32 AM
To: 'Stephanie Metzger'
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for your email.

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Thursday, July 29, 2010 10:25 AM
To: WhatWorks
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

What Works:

I appreciate all of your prompt replies to my extended series of emails. My questions/concerns below pertain to the Schwartz study conducted in 2005.

1. R.M. Schwartz authored, it appears, an article about "the history of Reading Recovery" and the "legacy of Marie Clay." Although the WWC did not consider this to be a study (and rightly so), it seems, right off the bat, that R.M. Schwartz is a "fan" of Reading Recovery. Again, as in previous emails, I raise concern that there is inherent bias on the part of the "researchers" who submitted their studies for consideration (i.e. Pinnell and company, Schwartz).

2. Students for the treatment (n=37) and comparison (n=37) were, as indicated, identified as "eligible" for Reading Recovery on the basis of their performance on 6 subtests of the Observation Survey and "teacher judgment." Teacher judgment was a factor in determining eligibility? Doesn't sound very objective to me. How were the students "randomly" assigned to the treatment and comparison groups? Was "teacher judgment" involved in the "random" assignment, or was the random assignment done through flip of the coin? How was the "random assignment" controlled?

3. Along the alphabetic domain, phonemic awareness, print awareness, letter identification and phonics construct effect sizes are reported by Schwartz. So, the effect on phonemic awareness was not statistically significant, based on the 10-item deletion task and the Yopp-Singer results. This is not surprising. Similarly, no statistically significant effect was found for letter identification (as per the WWC in contrast to the conclusion on the part of the researcher). The WWC confirmed that on two constructs (print awareness and "phonics") effect sizes were statistically significant, yet "print awareness" results were measured through the "concepts about print" subtest (my concerns with this subtest are expressed in a previous email) and "phonics" or phonetic skills were measured by the word recognition subtest from the Observation Survey. Again, the word recognition subtest does not measure whether a child has solid knowledge/awareness of the six syllable types (patterns) of English. One cannot rule out the fact that some words (e.g. the word "there") are merely "sight words." For the WWC to conclude that a word recognition subtest is a true measure of "phonics" tells me that this particular fact was either ignored, or that it was not understood by the WWC content experts and principal investigator.

4. It is not surprising, once again, that the children in the 1st round of Reading Recovery (treatment group) seemed to outperform the children who had to wait until 2nd round. While it appears, on its face, that the skills of the two groups were similar (matched) prior to the treatment, it is (again) clear that the children on the "wait list" simply "waited," getting only the regular classroom reading instruction. In other words, no control for "time on task" was indicated in the WWC's review

of the study. Again, any treatment that leads to more time engaged in learning is bound to yield positive effects. Could it be that the "time on task" factor alone, resulted in seemingly better results for the children in the treatment group?

5. As I stated in a previous email, the dictation and writing vocabulary subtests from the Observation Survey are not measures of general reading achievement!! They require auditory memory (dictation subtest) and strong spelling skills (both subtests).

Thank you for your time. Also know that I will be sending one more email with questions.

Stephanie Metzger

----- Forwarded Message -----

From: WhatWorks <WWorks@icfi.com>

To: Stephanie Metzger <(b)(6)>

Sent: Tue, July 27, 2010 1:18:03 PM

Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions. Please let us know when you have submitted all of your questions so that the QRT team can begin to work on a response.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]

Sent: Tuesday, July 27, 2010 11:54 AM

To: WhatWorks

Subject: Fw: What Works Clearinghouse (WWCPC 2230)

What Works:

I have not been able to peruse thhe WWC's conclusions regarding Reading Recovery for quite some time, but as I have stated below, have not forgotten. This

morning, I was able to read about the 1994 Pinnell, Lyons , DeFord, Bryk and Seltzer study. As you might suspect, I have serious concerns and questions. Immediately below I will try to convey them.

1. Of course, Pinnell and her fellow "researchers" are proponents of the Reading Recovery program and its methods. I do understand that the WWC does concern itself with "who conducted" the "research," as long as there are elements of experimental or quasi-experimental methods involved. I find this a serious problem, although I do understand that the number of studies meeting "the gold standard" are rare.

2. I find it interesting that the original "n" was 403 in 43 different schools, yet the actual number of students the WWC considered as it reviewed the study was a mere 79 students. Given that 31 students were randomly assigned to the "treatment" or intervention group and 48 were randomly assigned to "no intervention," and the actual study reported an original "n" of 403, it appears, on its face, that the WWC found it completely necessary to "dig" through the study to attempt to locate any element of random assignment at all. It is almost amusing that the authors of the study didn't find it necessary to compare students randomly assigned to Reading Recovery to those who received other types of interventions (Direct Instruction; Reading/Writing Group; Reading Successfully). True researchers understand that random assignment occurs across different treatment groups, and that effect sizes are calculated for each. I apologize, but the study should be thrown out on its face.

3. So, here we are. We have found at least some elements of the "study" that appeared sound, with a total "n" of 79 (31 who got the intervention and 48 who did not). It is now almost laughable that the 48 who did not receive the treatment, were children who received the "regular reading program." And what, pray tell, would that be? The children in the comparison group "did not receive any special instruction." Most educators know that anything is better than nothing at all. Time on task alone had to be a factor. Was time on task controlled? Have the WWC reviewers heard of the Matthew Effect cited in the literature? How are we to ever know whether the children in the control group were even exposed to daily, small group instruction, or that the time they spent engaged in "real reading" was comparable to the 30-minutes-a-day requirement of Reading Recovery? Was the "regular program" described in the study?

4. Once again, the dictation subtest (Hearing and Recording Sounds in Words) yielded an effect size of .65, but we all know that this subtest does, and never will,

measure "general reading achievement." It's incredible to me that the WWC would allow this effect size to even be considered in a rating along this construct. Also, why did the WWC not calculate a rating for the comprehension construct if the Gates MacGinitie Reading Test truly measures vocabulary and comprehension as described? Similarly, since I am familiar with the Woodcock Reading Mastery Test-R, I am wondering why there was no calculation of effect size for alphabets? Did the WWC simply decide to toss everything and just look at general reading achievement based on results across three assessments that all measure totally different things? This is unbelievable.

5. How did the study control for training that teachers had? The review alone demonstrates that the Reading Recovery teachers all had a certain amount of "training." Did the authors describe what kind of "training" the comparison group teachers had? Were they entry-level teachers? Hopefully, they were not paraprofessionals who delivered a regular reading program (no intervention) that I asked about in a previous email.

Thank you for your time.

Stephanie Metzger

----- Forwarded Message -----

From: Stephanie Metzger (b)(6)
To: WWorks@icfi.com
Sent: Tue, July 27, 2010 11:57:25 AM
Subject: Fw: What Works Clearinghouse (WWCPC 2230)

I apologize profusely. I forgot to attach! Here is the article.

----- Forwarded Message -----

From: Stephanie Metzger (b)(6)
To: WhatWorks <WWorks@icfi.com>
Sent: Tue, July 27, 2010 11:55:46 AM
Subject: Re: What Works Clearinghouse (WWCPC 2230)

WWC:

I thought I would pass along to you, an article appearing on the RRCNA (Reading Recovery Council of North America) website. It can be easily accessed online simply by going to the RRCNA homepage and clicking on the link about the National Center on Response to Intervention and how it has given Reading Recovery "high marks." I believe that you also ought to know that the RRCNA does seem to misrepresent the conclusions of the WWC. Flyers, articles, and other media publications from the RRCNA appear to exaggerate your actual conclusions.

I am reviewing the 1994 Pinnell study at this moment. I do want you to know that I appreciate your replies, and that I intend to continue to communicate with you as I read about the manner in which your ratings were reached. Our previous correspondence appears below. In the meantime, you may want to investigate the RRCNA website. Smoke and mirrors abound.

Stephanie

From: WhatWorks <WWorks@icfi.com>
To: Stephanie Metzger (b)(6)
Sent: Thu, June 24, 2010 4:50:07 PM
Subject: RE: What Works Clearinghouse (WWCPC 2230)

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received your email from earlier today, June 24, 2010. As mentioned previously, the QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Thursday, June 24, 2010 12:29 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2230)

What Works Clearinghouse:

Thank you very much for your quick reply. With regard to the Baenen study, I do want to say that I am encouraged that the WWC published its conclusions with "nr" (no rating) for several of the domains, but of course, wonder why "retention" was even considered as a reliable, valid measure of outcome.

I'd like to pose some additional concerns/questions, and appreciate the fact that the WWC will consider my questions/concerns jointly and is patient as I pose them, considering one study at a time.

What follow are my observations and concerns about the 1988 Pinnell study that involved 187 students in 14 schools in Columbus , Ohio .

1. I am encouraged that the WWC chose NOT to include 96 of the students whose classroom teachers were trained in Reading Recovery in its intervention rating, given that this circumstance falls outside the scope of the "standard" program.

2. Again, I have serious concerns about the use of Clay's Observations Survey as a valid assessment of reading outcomes. The WWC calculated effect sizes for the alphabetic and general reading achievement domains, given results of the Observation Survey. Here are my questions:

a) Does the "concepts about print" subtest really measure a print awareness construct as it falls under the alphabetic domain? Note that this subtest measures 22 "print concepts" deemed by its author to be valid measures of print awareness. Among items are those that require students to "show the front of the book," "show the back of the book," etc. How do items like these fall under the larger domain of "alphabetic" as defined by the National Reading Panel? The "effect size of .89, given this particular subtest, seems meaningless.

b) The "real" measure of print awareness, under the alphabetic domain, would be "letter identification," and the effect size was -.24 (not significant). It appears to me that the .89 effect size noted above (in section "a") masks the "real" measured outcome.

c) Also under the alphabetic domain is the phonics construct, and Clay's word recognition subtest (from the Observation Survey) was gauged by the WWC as an "acceptable" measure of this construct. This subtest is NOT a measure of phonetic knowledge. Note that there are six syllable patterns in the English language, and the majority of English words follow regular patterns (e.g. the word "interrupt" contains two closed syllables and one r-controlled syllable). The word recognition subtest measures, rather, "sight words" graded in terms of their difficulty levels. Sight words account for approximately 15% of the total number of English words...and all other words can be identified through their orthographic patterns. For the WWC to consider Clay's word recognition subtest as a true measure under the phonics construct is misleading, and ultimately, ignores what a "phonics construct" really is.

d) Within the scope of general reading achievement, two Observation Survey subtests were considered by the WWC (Dictation and Writing Vocabulary). I can hardly believe this. As a school psychologist, I have familiarized myself with these two subtests, and they are CLEARLY not measures of general reading achievement. The Dictation subtest is also called "Hearing and Recording Sounds in Words." The teacher dictates "a sentence" (e.g. "The bus will stop here to let me get on" and then the student must remember it verbatim, and try to write it (encode the words). Encoding means spelling. It is a different skill from reading, although poor readers are usually poor spellers as well. The subtest requires "writing" as already stated, on the part of the student, taxes short-term memory, and does not consider decoding or instant sight word recognition. It does not, in any way, concern itself with reading comprehension. It is hardly a measure of general reading achievement. Similarly, the writing vocabulary subtest, resulting in an effect size of .56, isn't a reading subtest either. The student is asked to "write as many words as he/she knows" within a specific amount of time. It, too, requires executive functions that are not really a part of "general reading." One is word generation. Poor writers often have difficulty "generating" words as they write. Generating while writing is an altogether different ability than comprehending while reading with speed, accuracy and expression. I have serious concerns with the WWC's conclusion that there are "positive effects" from participation in Reading Recovery in terms of general reading achievement based on these two subtests in the Pinnell study.

e) Finally, I would like to know why, when the n=88, only 82 students were included in the results of the CTBS Reading Comprehension subtest and 81 were included in the CTBS vocabulary subtest. What happened to the other students? As in my previous email, I'd like to reiterate that an "end-of-year" assessment may, or may not, be an adequate measure of the effects of Reading Recovery. If the 82/81 students were first round students, or if a portion of them were, what might have happened following the first round? Could they have been placed in another intervention? Attributing results of an end-of-year assessment to an intervention program without knowing "what else" might have happened in terms of intervention with students, is unacceptable.

Thank you for your time.

Sincerely,
Stephanie Metzger
Ohio

--- On Tue, 6/22/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2230)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 22, 2010, 12:03 PM

Dear Ms. Metzger,

Thank you for submitting your questions to the What Works Clearinghouse (WWC). We have received both of your emails from yesterday, June 21, 2010. Please note that you may submit your questions separately to the WWC; however, the Quality Review Team (QRT) will address them jointly. The QRT team will provide one response to all of your questions once we have received them. You may reply to this email to submit further questions.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Monday, June 21, 2010 4:08 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWCPC 2177)

I apologize. Here is the link to the abstract of a study of the Observation Survey.

<http://www.reading.org/Publish.aspx?page=/publications/journals/rrq/v41/i1/abstracts/rrq-41-1-denton.html&mode=redirect>

--- On **Mon, 6/21/10, Stephanie Metzger <steffmetzger@yahoo.com>** wrote:

From: Stephanie Metzger (b)(6)
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "WhatWorks " <WWorks@icfi.com>
Date: Monday, June 21, 2010, 4:55 PM

WWC:

I will provide my concerns about the WWC "standards" in a series of emails, if that is acceptable to you. They all surround the WWC's rating(s) of Reading Recovery.

First, I wish to ask how the 1997 Baenen (et.al.) study ever made its way to eventually meet the WWC's standards. It is clear that in reviewing this study, the WWC wisely chose to include only one of the cohorts (84 students who received the intervention and 84 who did not during the 90-91 school year...n=168) in its review of "effectiveness." However, "outcomes" in this study were "measured" in quite a dismal fashion. Clay's Observation Survey is hardly a valid assessment of "reading," and only 3 of the survey subtests were administered. I am attaching a summary of a study of Clay's Observation Survey for your consideration. Please keep in mind that although an "interrater agreement" study of the survey has been conducted in the past, the "text reading level" subtest is hardly a valid measure of "fluency." It does not consider speed at all (correct words per minute) but rather, makes use of a "running record," whereby teachers "tic off" words read correctly. This subtest is subjective, and hardly an adequate measure of speed, accuracy and expression...all described by the National Reading Panel as essential components of fluency.

Additionally, how could the WWC consider retention in first grade to be a reliable measure of "outcome"? I have been a school psychologist for decades, and it has always been clear to me (and to serious researchers) that retention decisions are subjective and fraught with problems. For one thing, when students receive Reading Recovery, its program's teachers are often involved in the retention decision-making process, and almost invariably, vote against retention for children they have served. Retention as an outcome? Really? You cannot be serious.

Finally, it is very clear that the state assessment in North Carolina was used as a third measure of "outcome." Your review does not describe whether the the 84 students who received the intervention (Reading Recovery) were in the first round, the second round, or were provided the intervention at the end of the year when a "slot" opened up and thus, may have received only a partial round. Clearly, if all 84 students were in the first round (and thus, a full round), who knows what "other" intervention may have been provided after the Reading Recovery teachers, themselves, were finished with the students. My experience has been that one has to "wait" until after the first round is over in order to change an intervention that is clearly not working for many, many students. Often, after the first round, children who are not "discontinued successfully" are then placed in small groups where teachers provide systematic instruction to

close a gap that has grown even wider during the 20 weeks they were enrolled in the program. Using an end-of-year assessment to measure outcome does not really mean anything, because it does not consider what else may have been tried for students who were struggling after they completed 20 weeks of the program.

I am amazed that the Observation Survey, retention decisions, and an end-of-year statewide assessment were touted in this study as reliable, valid measures of outcome of the Reading Recovery Program, and even more amazed that the WWC did not throw this study out on these bases alone.

Please respond.

Stephanie Metzger

--- On Tue, 6/8/10, WhatWorks <WWorks@icfi.com> wrote:

From: WhatWorks <WWorks@icfi.com>
Subject: RE: What Works Clearinghouse (WWCPC 2177)
To: "Stephanie Metzger" (b)(6)
Date: Tuesday, June 8, 2010, 12:10 PM

Dear Ms. Metzger,

The WWC takes seriously concerns raised about our reviews and has established a Quality Review Team to address any concerns. In keeping with the procedures for the Quality Review Team outlined in the [WWC Procedures and Standards Handbook](#), inquiries must be submitted in writing, pertain to a specific study or set of studies, identify the specific issue(s) in the review that the inquirer thinks are incorrect, and provide an explanation as to why the review may be incorrect.

Please submit your questions in response to this email at your convenience.

Thank you,

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/>.

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Saturday, June 05, 2010 10:16 AM
To: WhatWorks
Subject: Re: What Works Clearinghouse (WWCPC 2177)

I did not submit a study. I submitted a summary of a study, because the WWC failed to understand that the Iverson and Turner study that met the WWC's standards with reservations concluded that Reading Recovery would be more effective IF it was less rigid and its proponents willing to incorporate systematic phonics/phonemic awareness activities as part of its methodologies. I submitted the summary because I do not understand how the WWC concluded that Reading Recovery is an "effective" beginning reading program. Its theoretical underpinnings are the subject of controversy, it does not stay current with research, it relied on whole language methods, and the studies that met the WWC's standards (a total of 4) were all conducted by Reading Recovery advocates. I have found serious flaws in at least 3 of the studies, and it amazes me that the WWC reached its conclusions about "effectiveness" on these studies. I have asked repeatedly about this, was told that I would get a reply to my questions, and to date, have never received such. I believe, as do most serious reading educators (as well as respected researchers) that the WWC failed to live up to its promise to be a trusted entity examining effectiveness of educational programs. I will pose my questions again at a later date.

Stephanie Metzger

--- On **Mon, 5/24/10, WhatWorks <WhatWorks @icfi.com>** wrote:

From: WhatWorks <WhatWorks @icfi.com>
Subject: What Works Clearinghouse (WWCPC 2177)
To: [REDACTED]
Date: Monday, May 24, 2010, 5:52 PM

Hello,

Thank you for submitting a study for consideration of review by the What Works Clearinghouse (WWC). We successfully accessed the link that you sent.

We would like to tell you a bit more information about the timeline for reviews. As you may know, studies and interventions are not reviewed immediately upon request by the public, authors or developers. Rather, the WWC and IES determine the topics and interventions for review and the review teams proceed with the reviews in accordance with the topic protocols. We are currently unable to provide a timeline for the review process; IES is in the process of determining the timeline for upcoming reviews and has not released this information yet.

All publicly available information about future reviews is on our website. Please continue to check our website for updates at <http://ies.ed.gov/ncee/wwc/>.

There is a resource on our website called the WWC Procedures and Standards Handbook that provides detailed information about the WWC review process, including the evidence standards used when reviewing studies

(<http://ies.ed.gov/ncee/wwc/references/idocviewer/doc.aspx?docid=19&tocid=1>). Please note that the WWC reviews studies that are rigorous, outcome evaluations (such as randomized controlled trials or quasi-experimental designs).

We hope this information is helpful. If you have any questions about the WWC, please feel free

to respond to this email.

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: (b)(6)

Sent: Monday, May 24, 2010 9:55 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Evidence Standards, Reference ID Number:

1415747212

info@whatworks.ed.gov, this email was automatically sent through the Contact

link on the WWC website.

From: (b)(6)

Message: http://www.nrrf.org/rr_study_chapman.htm

From: What Works
Sent: Wednesday, March 20, 2013 10:52 AM
To: (b) (6)
Subject: What Works Clearinghouse (QR2010015)
Attachments: QR2010015.pdf

Dear Ms. Metzger,

Attached is a response to the questions you raised in your message regarding the What Works Clearinghouse findings for Reading Recovery®.

Thank you,

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/>.

From: Stephanie Metzger (b)(6)
Sent: Monday, April 15, 2013 12:32 PM
To: What Works
Subject: Re: What Works Clearinghouse (QR2010015)

Thank you for responding to my questions surrounding Reading Recovery. As a result, however, I have additional questions and comments. I will craft them and send them to you within the next month.

From: What Works <whatworks@mathematica-mpr.com>
To: (b)(6)
Sent: Wednesday, March 20, 2013 10:51 AM
Subject: What Works Clearinghouse (QR2010015)

Dear Ms. Metzger,

Attached is a response to the questions you raised in your message regarding the What Works Clearinghouse findings for Reading Recovery®.

Thank you,

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/>.

From: Stephanie Metzger (b)(6)
Sent: Tuesday, April 30, 2013 7:09 AM
To: What Works
Subject: Re: What Works Clearinghouse (QR2010015)
Attachments: reading-recovery20.pdf

Please see the attached summary of the existing research, including studies conducted outside the U.S.

From: Stephanie Metzger (b)(6)
To: What Works <whatworks@mathematica-mpr.com>
Sent: Monday, April 15, 2013 12:31 PM
Subject: Re: What Works Clearinghouse (QR2010015)

Thank you for responding to my questions surrounding Reading Recovery. As a result, however, I have additional questions and comments. I will craft them and send them to you within the next month.

From: What Works <whatworks@mathematica-mpr.com>
To: (b)(6)
Sent: Wednesday, March 20, 2013 10:51 AM
Subject: What Works Clearinghouse (QR2010015)

Dear Ms. Metzger,

Attached is a response to the questions you raised in your message regarding the What Works Clearinghouse findings for Reading Recovery®.

Thank you,

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/>.

From: WhatWorks
Sent: Tuesday, April 30, 2013 12:52 PM
To: (b)(6)
Subject: RE: What Works Clearinghouse (QR2010015) (WWCPC 2230)

Dear Ms. Metzger,

Thank you for your email and the attached article. Please let us know if you have specific questions for the WWC. You are welcome to reply to this email.

Thank you,
What Works Clearinghouse

From: Stephanie Metzger [<mailto:steffmetzger@yahoo.com>]
Sent: Tuesday, April 30, 2013 7:09 AM
To: What Works
Subject: Re: What Works Clearinghouse (QR2010015)

Please see the attached summary of the existing research, including studies conducted outside the U.S.

From: Stephanie Metzger <steffmetzger@yahoo.com>
To: What Works <whatworks@mathematica-mpr.com>
Sent: Monday, April 15, 2013 12:31 PM
Subject: Re: What Works Clearinghouse (QR2010015)

Thank you for responding to my questions surrounding Reading Recovery. As a result, however, I have additional questions and comments. I will craft them and send them to you within the next month.

From: What Works <whatworks@mathematica-mpr.com>
To: (b)(6) <steffmetzger@yahoo.com>
Sent: Wednesday, March 20, 2013 10:51 AM
Subject: What Works Clearinghouse (QR2010015)

Dear Ms. Metzger,

Attached is a response to the questions you raised in your message regarding the What Works Clearinghouse findings for Reading Recovery®.

Thank you,
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/>.

From: Stephanie Metzger (b)(6)
Sent: Thursday, May 02, 2013 9:02 PM
To: WhatWorks
Subject: Re: What Works Clearinghouse (QR2010015) (WWCPC 2230)I

Thank you for your prompt reply. I will prepare additional (specific) questions when I have time to do so (as a practicing school psychologist, I am very busy conducting multifactored evaluations....regrettably, for children who were once "recovered"...or should I say, "discontinued" from RR in grade one). While I appreciate your timely response, I find it disturbing that the WWC does not investigate the theoretical underpinnings of scientifically-based reading research with regard to beginning reading instruction.

From: WhatWorks <What.Works@icfi.com>
To: (b)(6)
Sent: Tuesday, April 30, 2013 12:52 PM
Subject: RE: What Works Clearinghouse (QR2010015) (WWCPC 2230)

Dear Ms. Metzger,

Thank you for your email and the attached article. Please let us know if you have specific questions for the WWC. You are welcome to reply to this email.

Thank you,
What Works Clearinghouse

From: Stephanie Metzger [mailto:(b)(6)]
Sent: Tuesday, April 30, 2013 7:09 AM
To: What Works
Subject: Re: What Works Clearinghouse (QR2010015)

Please see the attached summary of the existing research, including studies conducted outside the U.S.

From: Stephanie Metzger (b)(6)
To: What Works <whatworks@mathematica-mpr.com>
Sent: Monday, April 15, 2013 12:31 PM
Subject: Re: What Works Clearinghouse (QR2010015)

Thank you for responding to my questions surrounding Reading Recovery. As a result, however, I have additional questions and comments. I will craft them and send them to you within the next month.

From: What Works <whatworks@mathematica-mpr.com>

To: (b) (6)

Sent: Wednesday, March 20, 2013 10:51 AM

Subject: What Works Clearinghouse (QR2010015)

Dear Ms. Metzger,

Attached is a response to the questions you raised in your message regarding the What Works Clearinghouse findings for Reading Recovery®.

Thank you,

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/>.

What Works Clearinghouse WWC

A central and trusted source of scientific evidence for what works in education.

September 2, 2010

Ms. Stephanie Metzger

smetzger@mtrace.org

(b)(6)

Reference: QR201009

Dear Ms. Metzger:

Thank you for your inquiries concerning the WWC findings for *Reading Recovery*®. This letter is in response to your email of April 2010. I understand that you have submitted further inquiries. The WWC will respond to these inquiries at a later date.

In response to your April email, we conducted a 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. 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. These quality reviews are one of tools used to ensure that the standards established by the Institute of Educational Sciences (IES) are upheld on every review conducted by the WWC.

As you noted in your email, the WWC Intervention Report on *Reading Recovery*® (December 2008) concluded, “*Reading Recovery*® was found to have positive effects on alphabetic and general reading achievement and potentially positive effects on fluency and comprehension” based on four studies that meet WWC evidence standards and one study that meets WWC evidence standards with reservations. Your email mentioned your concerns with the four studies that meet WWC evidence standards. In conducting the quality review, we examined the four studies and the WWC ratings. The quality review concluded that the rating of “meets WWC evidence standards” is correct for each of these four studies. The quality review findings are described in greater detail in the remainder of this letter.

Under WWC standards, a study will receive a rating of “meets WWC evidence standards” if and only if the study is a well-designed randomized controlled trial that does not have problems with randomization or sample attrition. The WWC rating standards can be found in the WWC Handbook (pages 12-13), available at http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf. In conducting our review, we determined that each of the four studies cited in the Intervention Report meet these criteria.

What Works Clearinghouse WWC

A central and trusted source of scientific evidence for what works in education.

In your email, you ask specifically about the importance of ruling out bias on the part of the researchers. The WWC process is designed to carefully examine potential bias on the part of the researchers. For each study, two WWC reviewers, who are trained and certified in WWC standards, are assigned to independently examine the study design, outcomes, samples and attrition, and analysis methods. After they complete their review, they hold a reconciliation meeting with a senior WWC reviewer to discuss any differences between their reviews and any remaining issues about the study. During the reconciliation meeting, the reviewers discuss any potential concerns regarding bias on the part of the researchers. Ultimately, the study will receive the rating of “meets evidence standards” only if the authors clearly document that the evidence criteria have been met.

As you noted in your email, one of the assessments used in the studies was the Clay Observation Survey, which was developed by Dr. Marie Clay, who also developed *Reading Recovery*®. In evaluating the studies, the WWC reviewers examined the qualities of the Clay Observation Survey with respect to its validity, reliability, and consistency. In particular, the reviewers considered whether the survey was over-aligned with the intervention. For example, if the intervention involved practicing the assessment or reading materials that are used in the assessment, then the assessment would be over-aligned with the intervention. The WWC reviewers determined that there was no evidence of over-alignment. In addition, the Intervention Report cautions the reader about this potential concern (page 4).

Your other comments relate to the comparison groups against which the program was measured. The WWC does not have specific standards for comparison groups. In some cases, students who received an intervention will be compared to students who received only regular classroom instruction. In other cases, students who received an intervention will be compared to students who received an alternative intervention. WWC reviewers are trained to examine the comparison group and the “counterfactual condition” to determine their validity. For some studies with multiple comparisons, the WWC reports results from a subset of comparisons that meet WWC evidence standards. In the course of our quality review, we concluded that the comparison groups used for the results reported by the WWC Intervention Report are valid for measuring the impact of *Reading Recovery*®. The comparison groups for results reported by the WWC are as follows:

- In the Baenan et al. (1997) study, students who received *Reading Recovery*® were compared to students who qualified for *Reading Recovery*® but were randomly assigned to receive only regular services during the time period of the study.
- For the Pinnell et al. (1988) study, the WWC reviewers corresponded with the author in order to determine that the comparison group was composed of low-achieving students who did not have regular classroom teachers trained in *Reading Recovery*® and who were randomly assigned to an alternative compensatory program. These students received a series of drill activities conducted in small groups or individual sessions by a trained paraprofessional.

What Works Clearinghouse WWC

A central and trusted source of scientific evidence for what works in education.

- In the Pinnell et al. (1994) study, the comparison group included students who were randomly assigned to receive no special instruction but continued to participate in their regular reading program. These students were in the same schools as the intervention students in the study.
- In the Schwartz et al. (2005) study, the comparison group used for results reported by the WWC was composed of students who qualified for *Reading Recovery*®. These students were randomly assigned to receive *Reading Recovery*® in the second half of the year. These students served as comparisons only during the first part of the year when they received instruction in their regular classroom but did not receive the *Reading Recovery*® intervention.

As described above, in each of these four studies, students who received *Reading Recovery*® are compared to similar students who did not receive the program. Further descriptions of the comparison groups and other detailed information for these studies can be found in the technical appendices of the Beginning Reading Intervention Report, available at http://ies.ed.gov/ncee/wwc/pdf/wwc_reading_recovery_app_120208.pdf.

I hope this letter has addressed the concerns expressed in your April email.

Sincerely,

(b)(6)

Deborah Reed
WWC Quality Review Team

What Works Clearinghouse WWC

A central and trusted source of scientific evidence for what works in education.

September 2, 2010

Ms. Stephanie Metzger

smetzger@mtrace.org

(b)(6)

Reference: QR201009

Dear Ms. Metzger:

Thank you for your inquiries concerning the WWC findings for *Reading Recovery*[®]. This letter is in response to your email of April 2010. I understand that you have submitted further inquiries. The WWC will respond to these inquiries at a later date.

In response to your April email, we conducted a 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. 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. These quality reviews are one of tools used to ensure that the standards established by the Institute of Educational Sciences (IES) are upheld on every review conducted by the WWC.

As you noted in your email, the WWC Intervention Report on *Reading Recovery*[®] (December 2008) concluded, “*Reading Recovery*[®] was found to have positive effects on alphabetic and general reading achievement and potentially positive effects on fluency and comprehension” based on four studies that meet WWC evidence standards and one study that meets WWC evidence standards with reservations. Your email mentioned your concerns with the four studies that meet WWC evidence standards. In conducting the quality review, we examined the four studies and the WWC ratings. The quality review concluded that the rating of “meets WWC evidence standards” is correct for each of these four studies. The quality review findings are described in greater detail in the remainder of this letter.

Under WWC standards, a study will receive a rating of “meets WWC evidence standards” if and only if the study is a well-designed randomized controlled trial that does not have problems with randomization or sample attrition. The WWC rating standards can be found in the WWC Handbook (pages 12-13), available at http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf. In conducting our review, we determined that each of the four studies cited in the Intervention Report meet these criteria.

What Works Clearinghouse WWC

A central and trusted source of scientific evidence for what works in education.

In your email, you ask specifically about the importance of ruling out bias on the part of the researchers. The WWC process is designed to carefully examine potential bias on the part of the researchers. For each study, two WWC reviewers, who are trained and certified in WWC standards, are assigned to independently examine the study design, outcomes, samples and attrition, and analysis methods. After they complete their review, they hold a reconciliation meeting with a senior WWC reviewer to discuss any differences between their reviews and any remaining issues about the study. During the reconciliation meeting, the reviewers discuss any potential concerns regarding bias on the part of the researchers. Ultimately, the study will receive the rating of “meets evidence standards” only if the authors clearly document that the evidence criteria have been met.

As you noted in your email, one of the assessments used in the studies was the Clay Observation Survey, which was developed by Dr. Marie Clay, who also developed *Reading Recovery*®. In evaluating the studies, the WWC reviewers examined the qualities of the Clay Observation Survey with respect to its validity, reliability, and consistency. In particular, the reviewers considered whether the survey was over-aligned with the intervention. For example, if the intervention involved practicing the assessment or reading materials that are used in the assessment, then the assessment would be over-aligned with the intervention. The WWC reviewers determined that there was no evidence of over-alignment. In addition, the Intervention Report cautions the reader about this potential concern (page 4).

Your other comments relate to the comparison groups against which the program was measured. The WWC does not have specific standards for comparison groups. In some cases, students who received an intervention will be compared to students who received only regular classroom instruction. In other cases, students who received an intervention will be compared to students who received an alternative intervention. WWC reviewers are trained to examine the comparison group and the “counterfactual condition” to determine their validity. For some studies with multiple comparisons, the WWC reports results from a subset of comparisons that meet WWC evidence standards. In the course of our quality review, we concluded that the comparison groups used for the results reported by the WWC Intervention Report are valid for measuring the impact of *Reading Recovery*®. The comparison groups for results reported by the WWC are as follows:

- In the Baenan et al. (1997) study, students who received *Reading Recovery*® were compared to students who qualified for *Reading Recovery*® but were randomly assigned to receive only regular services during the time period of the study.
- For the Pinnell et al. (1988) study, the WWC reviewers corresponded with the author in order to determine that the comparison group was composed of low-achieving students who did not have regular classroom teachers trained in *Reading Recovery*® and who were randomly assigned to an alternative compensatory program. These students received a series of drill activities conducted in small groups or individual sessions by a trained paraprofessional.

What Works Clearinghouse WWC

A central and trusted source of scientific evidence for what works in education.

- In the Pinnell et al. (1994) study, the comparison group included students who were randomly assigned to receive no special instruction but continued to participate in their regular reading program. These students were in the same schools as the intervention students in the study.
- In the Schwartz et al. (2005) study, the comparison group used for results reported by the WWC was composed of students who qualified for *Reading Recovery*®. These students were randomly assigned to receive *Reading Recovery*® in the second half of the year. These students served as comparisons only during the first part of the year when they received instruction in their regular classroom but did not receive the *Reading Recovery*® intervention.

As described above, in each of these four studies, students who received *Reading Recovery*® are compared to similar students who did not receive the program. Further descriptions of the comparison groups and other detailed information for these studies can be found in the technical appendices of the Beginning Reading Intervention Report, available at http://ies.ed.gov/ncee/wwc/pdf/wwc_reading_recovery_app_120208.pdf.

I hope this letter has addressed the concerns expressed in your April email.

Sincerely,

(b)(6)

Deborah Reed
WWC Quality Review Team

From: WhatWorks
Sent: Thursday, September 02, 2010 2:57 PM
To: 'smetzger@mtrace.org'; 'Stephanie Metzger'
Subject: What Works Clearinghouse (WWCPC 2089)
Attachments: Response 201009.pdf

Dear Ms. Metzger,

Attached is a response to the questions you raised in your April 14 message to the What Works Clearinghouse (WWC).

Thank you,

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/>.

From: WhatWorks
Sent: Thursday, April 15, 2010 2:17 PM
To: 'smetzger@mtrace.org'
Subject: What Works Clearinghouse (WWCPC 2089)
Hello,

Thank you for contacting the What Works Clearinghouse (WWC). 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 was flawed, a revision will be published. These quality reviews are one of tools used to ensure that the standards established by the Institute of Educational Sciences (IES) are upheld on every review conducted by the What Works Clearinghouse.

Thank you,

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: smetzger@mtrace.org [mailto:smetzger@mtrace.org]

Sent: Wednesday, April 14, 2010 11:09 AM

To: info@whatworks.ed.gov

Subject: IES Website: Contact Us: Topic Areas, Beginning Reading Review,
Reference ID Number: 1852710946

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: smetzger@mtrace.org

Message: While I received the WWC's evidence standards, I continue to be perplexed about it's ratings for Reading Recovery. I have read the four studies that met the "standards." Two of them were conducted by Pinnell, who is introduced the program in the U.S. One of those studies compares the program to itself, with the only variable being the training of classroom teachers! The other study simply reveals that students in Reading Recovery performed better on the Observation Survey (created by Reading Recovery promoters) and Woodcock Johnson than students who received no intervention AT ALL. I just do not understand how the WWC can conclude that the program is effective based on the studies that met its standards. What is the WWC's view on the importance of ruling out bias on the part of the researchers? Please reply using my email shown above.