From: WhatWorks
Sent: Wednesday, July 07, 2010 12:27 PM
To: 'info@readnaturally.com'
Subject: What Works Clearinghouse (WWC 2259)

Dear Mr. Ihnot:

Thank you for contacting the What Works Clearinghouse (WWC). We received your voicemail from today, July 7, 2010.

Due to the need to maintain a written record of key decisions, the WWC can only respond to written questions. If you have questions about the WWC English Language Learners review, please submit them in writing and someone from the team will respond. You may respond 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: Karla Ramy [kramy@readnaturally.com] on behalf of info@readnaturally.com
Sent: Wednesday, July 07, 2010 12:53 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWC 2259)

WWC,
Please email the contact information for the lead reviewer for your English Language Learners section.

Thank you,
Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Maximize the benefits of your Read Naturally program. Attend one of our full-day training seminars near you!

Join Read Naturally on Facebook

From: WhatWorks [mailto:WWorks@icfi.com]
Sent: Wednesday, July 07, 2010 12:31 PM
To: info@readnaturally.com
Subject: What Works Clearinghouse (WWC 2259)

Dear Mr. Ihnot:

Thank you for contacting the What Works Clearinghouse (WWC). We received your voicemail from today, July 7, 2010.

Due to the need to maintain a written record of key decisions, the WWC can only respond to written questions. If you have questions about the WWC English Language Learners review, please submit them in writing and someone from the team will respond. You may respond 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: Wednesday, July 07, 2010 4:23 PM
To: 'info@readnaturally.com'
Subject: RE: What Works Clearinghouse (WWC 2259)

Dear Mr. Ihnot:

Due to the need to maintain a written record of key decisions, the WWC can only respond to written questions. If you have questions about the WWC English Language Learners review, please submit them in writing and someone from the team will respond. You may respond 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: Karla Ramy [mailto:kramy@readnaturally.com] **On Behalf Of** info@readnaturally.com
Sent: Wednesday, July 07, 2010 12:53 PM
To: WhatWorks
Subject: RE: What Works Clearinghouse (WWC 2259)

WWC,
Please email the contact information for the lead reviewer for your English Language Learners section.

Thank you,
Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Maximize the benefits of your Read Naturally program. Attend one of our full-day training seminars near you!

Join Read Naturally on Facebook

From: WhatWorks [mailto:WWorks@icfi.com]
Sent: Wednesday, July 07, 2010 12:31 PM
To: info@readnaturally.com
Subject: What Works Clearinghouse (WWC 2259)

Dear Mr. Ihnot:

Thank you for contacting the What Works Clearinghouse (WWC). We received your voicemail from today, July 7, 2010.

Due to the need to maintain a written record of key decisions, the WWC can only respond to written questions. If you have questions about the WWC English Language Learners review, please submit them in writing and someone from the team will respond. You may respond 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: Karla Ramy [kramy@readnaturally.com] on behalf of info@readnaturally.com
Sent: Friday, July 09, 2010 1:57 PM
To: What Works; WhatWorks
Subject: WWC 2259

It is very discouraging that WWC has selected another study to review that was not designed to evaluate Read Naturally.

The Kemp study was a study of SSR. It used Read Naturally passages but not the Read Naturally strategy. It was not a study of Read Naturally and should not be used to evaluate Read Naturally.

The Read Naturally group of students in the Kemp study achieved equivalent and significant growth in spite of the fact that Susan Kemp stated that the purpose of her study was not to evaluate Read Naturally. As a result, she did not adhere to several guidelines of the Read Naturally strategy. Notable discrepancies between the Kemp study and the intended use of Read Naturally include the following:

- Of the 157 students in Kemp's study, only 47 were at or below the 50th percentile in oral reading fluency based on the Hasbrouck/Tindal norms. The other 110 students were above the 50th percentile in fluency. Read Naturally is designed for students who are below the 50th percentile in fluency. Since 70 percent of the students in the study had fluency above the 50th percentile, Kemp could not have evaluated the effectiveness of the Read Naturally strategy and should not have used Read Naturally with these students.
- Kemp clearly stated on page 74 of her dissertation that Read Naturally was developed for at-risk readers:

  > "Another limitation [of the study] is that Read Naturally was not just used with at-risk readers. The Read Naturally program was developed to help students who are at-risk. Typically this would represent students who are scoring well below the 25th percentile. Students participating in this study did not fit the description of at-risk. Many scored well above the 25th percentile. Perhaps, if this study was only completed with at-risk students, results would have been different. In addition, because two of the participating schools were California Distinguished schools, the quality of instruction of the core program could have influenced results as well."
- Of the 47 students at or below the 50th percentile, only 28 were assigned to Read Naturally. The other 19 were assigned to SSR. The students assigned to Read Naturally gained an average of 16.8 words correct per minute (WCPM) in oral reading fluency, and the SSR group gained 15.2 WCPM, based on data from Dr. Kemp. Both groups' gains were significant.
- Read Naturally recommends a teacher-to-student ratio of 1 to 6. Kemp's study had a teacher-to-student ratio of 1 to 20. Consequently, the critical pass step would not have been completed by students and a teacher. Holding the students accountable at the pass step is critical for student success in Read Naturally. At the pass step the teacher checks whether the student can do the following:
  1. Read at the goal rate.
  2. Make three or fewer errors on the final timing.
  3. Read with good expression.

4. Answer all the comprehension questions correctly.

The teacher also checks the progress the student made in words correct per minute, from the cold timing of the story to the final timing. If the student fails any of the criteria, he is required to go back and master the failed criteria before he can move on to another passage.

In addition, the Kemp dissertation is flawed as a review of the Read Naturally strategy due to the following issues:

- The Kemp study excluded special education students from its analysis. These students, a target demographic for Read Naturally, would have likely performed poorly in SSR and performed well in Read Naturally.
- When comparing the Kemp study's Read Naturally group with the SSR group, the mean pretest scores in fluency for the native-speaking students in the "poor readers" category were substantially different. The mean score for native-speaking "poor readers" in the Read Naturally group was 40.5 WCPM, and the mean score for the native-speaking "poor readers" in the SSR group was 75.2 WCPM. Nonetheless, these disparate groups were compared.
- As stated above, the Kemp study imposed an intervention intended for students struggling with reading fluency on students who were not struggling with fluency. Seventy percent of the students in Kemp's study were above the 50th percentile in reading fluency.

It is also important to note that, in spite of the Kemp study's implementation flaws and comparison group design flaws, the Read Naturally group did make significant gains in fluency. The Read Naturally students averaged 1.5 WCPM per week. The Hasbrouck/Tindal data shows that the average fluency gain per week for third grade students is 1.1 WCPM. This data indicates that the students in the Read Naturally group were accelerating their progress. Kemp states in her study that both groups made significant gains in fluency.

Also, note that combining the Kemp study and the Denton study in the same ELL analysis produces two inappropriate reviews of Read Naturally. WWC does recognize that the Denton study modified Read Naturally, but the extent is not captured in its description. The other glaring problem with the Denton study is that the assessments were flawed. The Denton study was designed to evaluate Read Well. Consequently, the study used phonemic decoding as an assessment measure, and phonemic decoding is not part of the Read Naturally strategy. Also, the Denton study did not use fluency as an assessment measure and Read Naturally is regarded as the best fluency-building program (see attached analysis of the Denton study). http://www.readnaturally.com/approach/denton.htm

Please drop the Kemp study as a review of Read Naturally. If it remains on the WWC website it will mislead educators. I would appreciate the opportunity to have a phone conversation about the Kemp and Denton studies. Please contact me with the best time to call.

I look forward to discussing this issue with you.

Sincerely,

Tom Ihnot
CEO
Read Naturally, Inc.

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Give your intermediate-level readers a vocabulary boost with Take Aim! at Vocabulary, which includes theme-related, high-interest, nonfiction stories; repeated exposures to target words; and additional engaging activities.

From: WhatWorks
Sent: Wednesday, July 14, 2010 4:19 PM
To: 'info@readnaturally.com'
Subject: RE: WWC 2259

Dear Mr. Ihnot,

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

What Works Clearinghouse

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

From: Karla Ramy [mailto:kramy@readnaturally.com] **On Behalf Of** info@readnaturally.com
Sent: Friday, July 09, 2010 12:57 PM
To: What Works; WhatWorks
Subject: WWC 2259

It is very discouraging that WWC has selected another study to review that was not designed to evaluate Read Naturally.

The Kemp study was a study of SSR. It used Read Naturally passages but not the Read Naturally strategy. It was not a study of Read Naturally and should not be used to evaluate Read Naturally.

The Read Naturally group of students in the Kemp study achieved equivalent and significant growth in spite of the fact that Susan Kemp stated that the purpose of her study was not to evaluate Read Naturally. As a result, she did not adhere to several guidelines of the Read Naturally strategy. Notable discrepancies between the Kemp study and the intended use of Read Naturally include the following:

- Of the 157 students in Kemp's study, only 47 were at or below the 50th percentile in oral reading fluency based on the Hasbrouck/Tindal norms. The other 110 students were above the 50th percentile in fluency. Read Naturally is designed for students who are below the 50th percentile in fluency. Since 70 percent of the students in the study had fluency above the 50th percentile, Kemp could not have evaluated the effectiveness of the Read Naturally strategy and should not have used Read Naturally with these students.
- Kemp clearly stated on page 74 of her dissertation that Read Naturally was developed for at-risk readers:

  "Another limitation [of the study] is that Read Naturally was not just used with at-risk readers. The Read Naturally program was developed to help students who

are at-risk. Typically this would represent students who are scoring well below the 25th percentile. Students participating in this study did not fit the description of at-risk. Many scored well above the 25th percentile. Perhaps, if this study was only completed with at-risk students, results would have been different. In addition, because two of the participating schools were California Distinguished schools, the quality of instruction of the core program could have influenced results as well."

- Of the 47 students at or below the 50th percentile, only 28 were assigned to Read Naturally. The other 19 were assigned to SSR. The students assigned to Read Naturally gained an average of 16.8 words correct per minute (WCPM) in oral reading fluency, and the SSR group gained 15.2 WCPM, based on data from Dr. Kemp. Both groups' gains were significant.
- Read Naturally recommends a teacher-to-student ratio of 1 to 6. Kemp's study had a teacher-to-student ratio of 1 to 20. Consequently, the critical pass step would not have been completed by students and a teacher. Holding the students accountable at the pass step is critical for student success in Read Naturally. At the pass step the teacher checks whether the student can do the following:
    1. Read at the goal rate.
    2. Make three or fewer errors on the final timing.
    3. Read with good expression.
    4. Answer all the comprehension questions correctly.

    The teacher also checks the progress the student made in words correct per minute, from the cold timing of the story to the final timing. If the student fails any of the criteria, he is required to go back and master the failed criteria before he can move on to another passage.

In addition, the Kemp dissertation is flawed as a review of the Read Naturally strategy due to the following issues:

- The Kemp study excluded special education students from its analysis. These students, a target demographic for Read Naturally, would have likely performed poorly in SSR and performed well in Read Naturally.
- When comparing the Kemp study's Read Naturally group with the SSR group, the mean pretest scores in fluency for the native-speaking students in the "poor readers" category were substantially different. The mean score for native-speaking "poor readers" in the Read Naturally group was 40.5 WCPM, and the mean score for the native-speaking "poor readers" in the SSR group was 75.2 WCPM. Nonetheless, these disparate groups were compared.
- As stated above, the Kemp study imposed an intervention intended for students struggling with reading fluency on students who were not struggling with fluency. Seventy percent of the students in Kemp's study were above the 50th percentile in reading fluency.

It is also important to note that, in spite of the Kemp study's implementation flaws and comparison group design flaws, the Read Naturally group did make significant gains in fluency. The Read Naturally students averaged 1.5 WCPM per week. The Hasbrouck/Tindal data shows that the average fluency gain per week for third grade students is 1.1 WCPM. This data indicates that the students in the Read Naturally group were accelerating their progress. Kemp states in her study that both groups made significant gains in fluency.

Also, note that combining the Kemp study and the Denton study in the same ELL analysis produces two inappropriate reviews of Read Naturally. WWC does recognize that the Denton study modified Read Naturally, but the extent is not captured in its description. The other glaring problem with the Denton study is that the assessments were flawed. The Denton study was designed to evaluate Read Well. Consequently, the study used phonemic decoding as an

assessment measure, and phonemic decoding is not part of the Read Naturally strategy. Also, the Denton study did not use fluency as an assessment measure and Read Naturally is regarded as the best fluency-building program (see attached analysis of the Denton study). http://www.readnaturally.com/approach/denton.htm

Please drop the Kemp study as a review of Read Naturally. If it remains on the WWC website it will mislead educators. I would appreciate the opportunity to have a phone conversation about the Kemp and Denton studies. Please contact me with the best time to call.

I look forward to discussing this issue with you.

Sincerely,

Tom Ihnot
CEO
Read Naturally, Inc.

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Give your intermediate-level readers a vocabulary boost with Take Aim! at Vocabulary, which includes theme-related, high-interest, nonfiction stories; repeated exposures to target words; and additional engaging activities.

From: Karla Ramy [kramy@readnaturally.com] on behalf of info@readnaturally.com
Sent: Thursday, July 15, 2010 2:21 PM
To: WhatWorks
Subject: Tom Ihnot

Follow Up Flag: Follow up
Flag Status: Completed

Attachments: Flaws in the Denton Study Reviewed by What Works Clearinghouse.docx; Flaws in the Kemp Study Reviewed by WWC.docx

WWC,

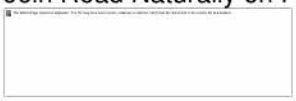
Since you are preparing a response, I am sending you a REVISED analysis of the Kemp and Denton studies.

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Maximize the benefits of your Read Naturally program. Attend one of our full-day training seminars near you!

Join Read Naturally on Facebook



From: WhatWorks
Sent: Friday, July 16, 2010 11:30 AM
To: 'info@readnaturally.com'
Subject: RE: Tom Ihnot

Dear Mr. Ihnot:

We have received your email and attachments.

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: Karla Ramy [mailto:kramy@readnaturally.com] **On Behalf Of** info@readnaturally.com
Sent: Thursday, July 15, 2010 2:21 PM
To: WhatWorks
Subject: Tom Ihnot

WWC,

Since you are preparing a response, I am sending you a REVISED analysis of the Kemp and Denton studies.

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Maximize the benefits of your Read Naturally program. Attend one of our full-day training seminars near you!

Join Read Naturally on Facebook

February 26, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013

Dear Mr. Ihnot:

Thank you for your inquiries concerning the What Works Clearinghouse (WWC) reviews of four studies of Read Naturally®—Kemp (2006), Denton et al. (2004), Hancock (2002), and Chenault et al. (2006)—and for your email regarding my response to Dr. Kevin Feldman's inquiry about the WWC reviews of studies of Read Naturally®. In response to your concerns, we conducted an independent quality review. The WWC quality review team responds to concerns raised by study authors, curriculum developers, or other relevant parties about WWC reviews published on our website. When a quality review is conducted, a researcher who was not involved in the initial reviews undertakes an independent assessment of the studies in question. The researcher also investigates the procedures used and decisions made during the original reviews of the studies. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the WWC.

The quality review considered the concerns from your emails regarding study inclusion, implementation, student sample, and outcomes.

Study Inclusion

The WWC followed protocol in choosing to review these four studies. The WWC screens studies based on a number of factors, including relevancy and methodology criteria (as described in the WWC Procedures and Standards Handbook, pages 8–10, available at http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19). The WWC does not screen based on whether the author(s) explicitly intended the study as an evaluation of an intervention, but rather whether the study presents a primary analysis of the effect of an intervention. This screening process allows for the review of a wide range of relevant and methodologically sound studies.

Thank you for bringing to my attention the studies by Mesa (2004) and Tucker (2010). The Tucker (2010) study will be included in our literature search for updating the Intervention Report

for Read Naturally®, reviewed under the Beginning Reading Evidence Review Protocol. In addition, as you note, the Mesa (2004) study was included in the original review of Read Naturally®, but will be reviewed again in an updated review.

Implementation

The WWC followed protocol in the manner in which the four studies are described in WWC publications. The WWC Handbook says, "The WWC makes no adjustments or corrections for variations in implementation of the intervention; however, if a study meets standards and is included in an intervention report, descriptions of implementation are provided in the report appendices to provide context for the findings" (page 16). This approach is appropriate because there is no standard metric for fidelity to intervention design. Thus, the WWC includes studies with variation in fidelity and does not evaluate implementation fidelity. The quality review team verified that variations in implementation that are noted in the four studies and that may affect the interpretation of findings were properly included in the WWC publications.

Regarding the Kemp (2006) study, your emails raised a concern about implementation of the Read Naturally® program. The quality review noted that the study does not provide any indication of deviation from program design. Specifically, the study states, "… it could be concluded that all teachers implemented the Read Naturally® program as prescribed" (page 40). Based on this information, the quality review concluded that there was no evidence that deviations in implementation should have been noted in the WWC Intervention Report. Regarding your specific concern about the teacher-to-student ratio, the study does not note any deviation from the Read Naturally® program in this area and does not provide any information on the ratio.

You also expressed concern over deviations in the implementation of Read Naturally® in the Denton et al. (2004) study. The quality review found that the inclusion of additional activities was noted in the study and reported in the WWC Intervention Report (see footnote 7 on page 3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). The other potential deviations that you noted (accountability to pass criteria, use of audio tapes, and the number of sessions per week) were not identified in the study and are therefore not noted in the WWC publication.

Regarding your concerns that the Read Naturally® strategy was not implemented fully for the Hancock (2002) study, the quality review found that WWC procedures were followed in the WWC Intervention Report. The WWC contacted Hancock regarding the implementation of Read Naturally® in the 2002 study. Her response indicated that the study excluded Read Naturally®'s pre-reading vocabulary instruction component and the placement system to individualize instruction. The WWC Intervention Report was revised to note these deviations in implementation (see footnote 4 on page 2, available at http://ies.ed.gov/ncee/wwc/InterventionReport.aspx?sid=407). Other potential concerns with implementation described in your email were not reported by the author and were not included in the revised WWC Intervention Report.

The quality review also investigated concerns that you raised in prior emails. The email from June 2012 regarding Dr. Feldman's questions was addressed in my letter to Dr. Feldman; I will not

repeat the quality review findings here. Your email from May 2012 included some text that was reported to be a quote from Hancock's email to the WWC, but the quoted statements were not part of Hancock's email text. The misreported quote is also found on the Read Naturally® website at http://www.readnaturally.com/approach/hancock.htm.

Regarding the Chenault et al. (2006) study, you expressed concern with the implementation of Read Naturally® related to the selection of students and the limited number of sessions. The quality review found that the WWC Intervention Report appropriately notes both of these factors (see pages 2–3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=409).

Student Sample

You also raised a concern that the Kemp (2006) study included a large share of students who were above the 50th percentile in reading fluency who were not "at-risk" readers. The quality review verified that the WWC Intervention Report accurately states that the study population for the results reported by the WWC was third-grade English language learner (ELL) students (page 2, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). Furthermore, the quality review noted that Read Naturally® program materials state that the program "… products help students of all ages and abilities—whether special education, ELL, Title I, or mainstream—to improve their reading skills, gain confidence, and transform their attitudes toward school and learning" (available at http://www.readnaturally.com/products/default.htm).

Finally, regarding the differences in pre-test performance among native-speaking students in the treatment and control groups, the quality review found that these concerns are not relevant to this WWC publication, which reports on findings only for English language learner students. Furthermore, the quality review concluded that the WWC appropriately restricted the reported findings to impacts assessed relative to the comparison group, and thus did not include achievement gains in oral reading fluency.

Outcomes

Regarding the Denton et al. (2004) study, you expressed concern that the study did not assess fluency. The quality review noted that the WWC Intervention Report clearly states that the study assessed reading achievement (page 4, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408), an outcome of relevance to the WWC ELL topic area. The quality review noted that the Read Naturally® program materials state there are several components of the program beyond fluency that are related to reading achievement: "Our industry-leading programs develop and support the five essential components of reading, identified by the National Reading Panel: phonemic awareness, phonics, fluency, vocabulary, and comprehension" (see "The Read Naturally Strategy," available at http://www.readnaturally.com/approach/default.htm).

I hope that this letter has addressed your concerns. You requested that we meet to discuss the WWC reviews of Read Naturally®. However, WWC policy requires all correspondence in writing. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov.

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

cc: Joy Lesnick, IES

Sent:26 Feb 2013 20:09:42 +0000
To:'kramy@readnaturally.com';'info@readnaturally.com'
Subject:What Works Clearinghouse (QR2010013)
Attachments:QRT2010013.pdf
Dear Mr. Ihnot:

Attached is a response to the questions you raised in your inquiries concerning the What Works Clearinghouse reviews of studies of Read Naturally®.

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/.

Flaws in the Denton Study Reviewed by What Works Clearinghouse

The Carolyn Denton study, "The Effects of Two Tutoring Programs on the English Reading Development of Spanish-English Bilingual Students" was published in the Elementary School Journal in 2004. **This study's assessment tools and process were incongruous with the objectives of the Read Naturally strategy.**

Denton Study Assessments Flawed

The Denton study used two programs in its study: Read Naturally and Read Well. Read Naturally is regarded as the best fluency building program on the market that also improves comprehension and vocabulary, yet **the Denton study did not assess fluency**.

The Denton study did assess phonemic decoding, which is not part of the Read Naturally strategy. Using phonemic decoding as an assessment measure benefitted Read Well and negatively impacted Read Naturally. Clearly, the Denton study was designed to measure the effects of Read Well.

One of the authors of the study, **Dr. Jan Hasbrouck**, acknowledges these flaws. Hasbrouck states: "The measures used to evaluate the effectiveness of the instruction did not include measures of fluency. Since Read Naturally is an intervention targeted at fluency, *the measures used did **not** match the specific purpose and intent of the Read Naturally instruction* given to the students in that group."

Denton Study Process Flawed

Even f the Denton study had assessed fluency, its process would have produced an inaccurate evaluation of the Read Naturally strategy. The Denton study did not follow Read Naturally's recommended process. Specifically:

- The Denton study's inclusion of oral discussion sessions of vocabulary and comprehension in the process would have **reduced the time spent reading**.

  The Denton study's inclusion of vocabulary activities such as flash cards and pre-reading in the process would also have **reduced the time spent reading.**

- The students in the Denton study were **not held accountable to Read Naturally's four criteria to pass** (reach goal rate, make three or fewer errors, read with good expression, and answer all of the questions correctly). Students were not required to go back and master any of the criteria they failed on the first pass attempt. Achieving mastery in all four of the pass criteria is a critical element to success in the Read Naturally process.

- The Denton study's **use of audio tapes was inconsistent**. Variability of reading rates would have resulted from having tutors—volunteer undergraduate students—read to the students. Rates on the audio tapes increase by a specific percentage between read alongs, which tutors would not have been able to replicate.

- The students in the study only had 22 sessions over a 10-week period. Read Naturally recommends a minimum of three sessions per week to be successful.

Clearly, the results of the Denton study cannot be used to accurately evaluate the effectiveness of Read Naturally. WWC's review of this study as though it does evaluate Read Naturally provides the public with inaccurate information.

Flaws in the Kemp Study Reviewed by What Works Clearinghouse

Susan Kemp conducted a study that used Read Naturally passages but not the Read Naturally strategy. It was a study of SSR (Sustained Silent Reading). Susan Kemp stated that the purpose of her study was not to evaluate Read Naturally. Because the Kemp study did not intend to evaluate Read Naturally, its results cannot be interpreted as evidence of the program's effectiveness.

Notable discrepancies between the Kemp study and the intended use of Read Naturally include the following:

- **Most students in the Kemp study were above the 50th percentile in reading fluency.** Of the 157 students in the study, only 47 were at or below the 50th percentile in oral reading fluency as determined by the Hasbrouck/Tindal oral reading fluency norms. The other 110 students—70 percent of the study participants—were above the 50th percentile in fluency. Read Naturally is designed for students below the 50th percentile in fluency. Seventy percent of those students studied were not a group for whom a Read Naturally intervention is intended.

- **Susan Kemp acknowledges that the study participants did not need Read Naturally.** Kemp states on page 74 of her dissertation that Read Naturally is intended for at-risk readers, which is not the population she studied:

  > "Another limitation [of the study] is that Read Naturally was not just used with at-risk readers. The Read Naturally program was developed to help students who are at-risk. Typically this would represent students who are scoring well below the 25th percentile. Students participating in this study did not fit the description of at-risk. Many scored well above the 25th percentile. Perhaps, if this study was only completed with at-risk students, results would have been different. In addition, because two of the participating schools were California Distinguished schools, the quality of instruction of the core program could have influenced results as well."

- **The critical pass step was not implemented correctly.** Read Naturally recommends a teacher-to-student ratio of 1 to 6. Kemp's study had a teacher-to-student ratio of 1 to 20. Consequently, the pass step of the Read Naturally process could not be implemented with fidelity. Holding the students accountable at the pass step is critical for student success in Read Naturally.

In addition, the Kemp study is flawed as a review of the Read Naturally strategy due to the following issues:

- **Special Education students were excluded from the study.** These students, a target demographic for Read Naturally, would have likely performed poorly in SSR and performed well in Read Naturally.

- **Disparate groups were compared.** When comparing the Kemp study's Read Naturally group with the SSR group, the mean pretest scores in fluency for the native-speaking students in the "poor readers" category were substantially different. The mean score for native-speaking "poor readers" in the Read Naturally group was 40.5 WCPM, and the mean score for native-speaking "poor readers" in the SSR group was 75.2 WCPM. Nonetheless, these disparate groups were compared.

In spite of the Kemp study's design flaws, the group using Read Naturally passages did make significant gains in fluency. These students averaged a gain of 1.5 WCPM (Words Correct per Minute) each week. The Hasbrouck/Tindal data shows that the average fluency gain per week for third grade students is 1.1 WCPM. This data indicates that the students in the Read Naturally group were accelerating their progress.

Nonetheless, because the Kemp study did not follow the Read Naturally strategy and was not intended as an evaluation of Read Naturally, WWC should not have reviewed it as though it was. Any conclusions drawn about Read Naturally's effectiveness on the basis of this study are therefore inaccurate. Furthermore, WWC combines the Kemp study and the Denton study in the same ELL analysis. The Denton study was designed to evaluate Read Well, not Read Naturally. This study used phonemic decoding, which is not part of the Read Naturally strategy, as an assessment measure. The Denton study did not use fluency as an assessment measure, even though Read Naturally is a fluency-building program.

From: WhatWorks
Sent: Tuesday, July 27, 2010 11:33 AM
To: 'info@readnaturally.com'
Subject: re: Mark Dynarski - Director, What Works Clearinghouse (WWC 2304)

Dear Mr. Ihnot,

Thank you for contacting the What Works Clearinghouse (WWC). The WWC Quality Review Team is reviewing your email and will prepare a response.

What Works Clearinghouse

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

From: Karla Ramy [mailto:kramy@readnaturally.com] **On Behalf Of** info@readnaturally.com
Sent: Monday, July 26, 2010 4:20 PM
To: What Works
Subject: Mark Dynarski - Director, What Works Clearinghouse

Mark Dynarski
Director, What Works Clearinghouse,

When determining whether or not to use a particular study as a review of a reading program, WWC needs to include criteria such as the purpose of the study and the nature of the study's implementation. Otherwise, WWC will continue to review programs based on studies that are not designed to examine the program and do not follow the program's procedures.

I have previously sent you analyses of the Hancock, Denton, and Kemp studies (see attached). These analyses clearly demonstrate that the authors of the studies did not intend to evaluate Read Naturally and did not follow the implementation procedures of the Read Naturally strategy. As a consequence, WWC has misinformed educators about Read Naturally.

The Chenault study is another gross injustice to Read Naturally and the educators who use WWC's website. Belle Montgomery Chenault clearly states on pages 53 and 54 of her dissertation that she did not intend to evaluate Read Naturally:

> *This study was not intended as an examination of the Read Naturally program. The lack of demonstrable standard score gains on the GORT Rate subtest for the Read Naturally group following this short term of training should not be construed as indicating that this instructional program is not effective in improving reading fluency. These children were not selected for their reading fluency problems – they were selected for their persisting writing problems.*

In addition, I cannot believe WWC does not have a length-of-study criteria. In the Chenault study ten sessions of twenty-five minutes (two weeks) is completely inadequate to evaluate a program. Students take two weeks to learn the steps and behavior to be successful in the Read Naturally strategy.

I strongly urge you not to post the Chenault study as a review of Read Naturally. It was designed to evaluate Pay Attention and Writing Composition.

I will send a more detailed analysis of WWC's use of Chenault as an evaluation of Read Naturally after I have had a chance to talk to the study's authors. I would appreciate the opportunity to discuss this issue with the lead reviewer of the learning disabilities section. Please contact me at 651-286-8721.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Flaws in the Denton Study Reviewed by What Works Clearinghouse

The Carolyn Denton study, "The Effects of Two Tutoring Programs on the English Reading Development of Spanish-English Bilingual Students" was published in the Elementary School Journal in 2004. **This study's assessment tools and process were incongruous with the objectives of the Read Naturally strategy.**

Denton Study Assessments Flawed

The Denton study used two programs in its study: Read Naturally and Read Well. Read Naturally is regarded as the best fluency building program on the market that also improves comprehension and vocabulary, yet **the Denton study did not assess fluency**.

The Denton study did assess phonemic decoding, which is not part of the Read Naturally strategy. Using phonemic decoding as an assessment measure benefitted Read Well and negatively impacted Read Naturally. Clearly, the Denton study was designed to measure the effects of Read Well.

One of the authors of the study, **Dr. Jan Hasbrouck**, acknowledges these flaws. Hasbrouck states: "The measures used to evaluate the effectiveness of the instruction did not include measures of fluency. Since Read Naturally is an intervention targeted at fluency, *the measures used did **not** match the specific purpose and intent of the Read Naturally instruction* given to the students in that group."

Denton Study Process Flawed

Even f the Denton study had assessed fluency, its process would have produced an inaccurate evaluation of the Read Naturally strategy. The Denton study did not follow Read Naturally's recommended process. Specifically:

- The Denton study's inclusion of oral discussion sessions of vocabulary and comprehension in the process would have **reduced the time spent reading**.

  The Denton study's inclusion of vocabulary activities such as flash cards and pre-reading in the process would also have **reduced the time spent reading.**

- The students in the Denton study were **not held accountable to Read Naturally's four criteria to pass** (reach goal rate, make three or fewer errors, read with good expression, and answer all of the questions correctly). Students were not required to go back and master any of the criteria they failed on the first pass attempt. Achieving mastery in all four of the pass criteria is a critical element to success in the Read Naturally process.

- The Denton study's **use of audio tapes was inconsistent**. Variability of reading rates would have resulted from having tutors—volunteer undergraduate students—read to the students. Rates on the audio tapes increase by a specific percentage between read alongs, which tutors would not have been able to replicate.

- The students in the study only had 22 sessions over a 10-week period. Read Naturally recommends a minimum of three sessions per week to be successful.

Clearly, the results of the Denton study cannot be used to accurately evaluate the effectiveness of Read Naturally. WWC's review of this study as though it does evaluate Read Naturally provides the public with inaccurate information.

Flaws in the Hancock Study Reviewed by What Works Clearinghouse

The Carrie Hancock dissertation study, "Accelerating Reading Trajectories: The Effects of Dynamic Research Based Instructions" **used Read Naturally passages but did not follow Read Naturally recommended process.** In fact, Carrie Hancock sent an e-mail message to What Works Clearinghouse (WWC) to explain to them that they should not publish a review of her study as Read Naturally. In her e-mail Carrie stated, "While I used Read Naturally materials, I did **NOT** fully implement the Read Naturally strategy and my study was **NOT** intended to evaluate the Read Naturally strategy. Rather, the purpose was to determine the impact of ongoing supplemental fluency practice on second grade students rates of learning to read."

The response from Becki Herman, WWC Project Director to Hancock's e-mail was that WWC does not consider the purpose of the study and does not consider implementation.

The Hancock study did **NOT** intend to use the Read Naturally strategy. The following are **some of the deviations the Hancock study made from the Read Naturally strategy**:

- The purpose of the Hancock study was to measure the rates of learning to read for second graders. It was **not designed to use the Read Naturally strategy with struggling readers and measure the effectiveness of Read Naturally**.
- The Hancock study **did not use the Read Naturally placement system** to ensure students started at a level where they would be challenged and be successful.
- It did not set goals based on an initial Read Naturally placement.
- **Goals were not individualized** based on the Read Naturally placement system. The Hancock study had the same goal of 100 WCPM for every student.
- The Hancock process of every student doing three read alongs and five one minute practice timings and then trying to achieve 100 WCPM was very inefficient. It would have been a punitive process for struggling students who could not achieve the goal in three read alongs and five practice reads and a boring process for students who could achieve the goal without eight readings of the passage.
- There was no adjustment of goal and level based on cold and final timings based on the Read Naturally goal and level adjustment guidelines.
- There was **no prediction step, no key words step, and no retell step**.
- The Hancock study **did not have the Read Naturally recommended four criteria pass step** (reach individualized goal rate, three or fewer errors, read with good expression and answer all the questions correctly. Students were not required to go back and master the expression or questions criteria if they fail on the pass step.
- **A group size of 10 to 12** would have made it impossible for a teacher assistant to get to students in a timely manner for cold and final timings. Students would have been sitting around waiting.
- Not following Read Naturally's placement, initial goal setting, adjusting goals and levels guidelines and progress monitoring procedures would have destroyed the intrinsic motivation that is a hallmark of the Read Naturally strategy.

Flaws in the Kemp Study Reviewed by What Works Clearinghouse

Susan Kemp conducted a study that used Read Naturally passages but not the Read Naturally strategy. It was a study of SSR (Sustained Silent Reading). Susan Kemp stated that the purpose of her study was not to evaluate Read Naturally. Because the Kemp study did not intend to evaluate Read Naturally, its results cannot be interpreted as evidence of the program's effectiveness.

Notable discrepancies between the Kemp study and the intended use of Read Naturally include the following:

- **Most students in the Kemp study were above the 50th percentile in reading fluency.** Of the 157 students in the study, only 47 were at or below the 50th percentile in oral reading fluency as determined by the Hasbrouck/Tindal oral reading fluency norms. The other 110 students—70 percent of the study participants—were above the 50th percentile in fluency. Read Naturally is designed for students below the 50th percentile in fluency. Seventy percent of those students studied were not a group for whom a Read Naturally intervention is intended.

- **Susan Kemp acknowledges that the study participants did not need Read Naturally.** Kemp states on page 74 of her dissertation that Read Naturally is intended for at-risk readers, which is not the population she studied:

  > "Another limitation [of the study] is that Read Naturally was not just used with at-risk readers. The Read Naturally program was developed to help students who are at-risk. Typically this would represent students who are scoring well below the 25th percentile. Students participating in this study did not fit the description of at-risk. Many scored well above the 25th percentile. Perhaps, if this study was only completed with at-risk students, results would have been different. In addition, because two of the participating schools were California Distinguished schools, the quality of instruction of the core program could have influenced results as well."

- **The critical pass step was not implemented correctly.** Read Naturally recommends a teacher-to-student ratio of 1 to 6. Kemp's study had a teacher-to-student ratio of 1 to 20. Consequently, the pass step of the Read Naturally process could not be implemented with fidelity. Holding the students accountable at the pass step is critical for student success in Read Naturally.

In addition, the Kemp study is flawed as a review of the Read Naturally strategy due to the following issues:

- **Special Education students were excluded from the study.** These students, a target demographic for Read Naturally, would have likely performed poorly in SSR and performed well in Read Naturally.

- **Disparate groups were compared.** When comparing the Kemp study's Read Naturally group with the SSR group, the mean pretest scores in fluency for the native-speaking students in the "poor readers" category were substantially different. The mean score for native-speaking "poor readers" in the Read Naturally group was 40.5 WCPM, and the mean score for native-speaking "poor readers" in the SSR group was 75.2 WCPM. Nonetheless, these disparate groups were compared.

In spite of the Kemp study's design flaws, the group using Read Naturally passages did make significant gains in fluency. These students averaged a gain of 1.5 WCPM (Words Correct per Minute) each week. The Hasbrouck/Tindal data shows that the average fluency gain per week for third grade students is 1.1 WCPM. This data indicates that the students in the Read Naturally group were accelerating their progress.

Nonetheless, because the Kemp study did not follow the Read Naturally strategy and was not intended as an evaluation of Read Naturally, WWC should not have reviewed it as though it was. Any conclusions drawn about Read Naturally's effectiveness on the basis of this study are therefore inaccurate. Furthermore, WWC combines the Kemp study and the Denton study in the same ELL analysis. The Denton study was designed to evaluate Read Well, not Read Naturally. This study used phonemic decoding, which is not part of the Read Naturally strategy, as an assessment measure. The Denton study did not use fluency as an assessment measure, even though Read Naturally is a fluency-building program.

info@readnaturally.com

From:info@readnaturally.com
Sent:28 Jul 2010 13:47:26 -0500
To:What Works
Subject:Robert G St. Pierre
Attachments:Flaws in the Hancock Study Reviewed.doc, Flaws in the Kemp Study Reviewed by WWC.DOCX, Flaws in the Denton Study Reviewed by What Works Clearinghouse.docx

Robert

When determining whether or not to use a particular study as a review of a reading program, WWC needs to include criteria such as the purpose of the study and the nature of the study's implementation. Otherwise, WWC will continue to review programs based on studies that are not designed to examine the program and do not follow the program's procedures.

I have previously sent WWC an analysis of the Hancock, Denton, and Kemp studies (see attached). These reviews clearly demonstrate that the authors of the studies did not intend to evaluate Read Naturally and did not follow the implementation procedures of the Read Naturally strategy. As a consequence, WWC has misinformed educators about Read Naturally.

The Chenault study is another gross injustice to Read Naturally and the educators who use WWC's website. Belle Montgomery Chenault clearly states on pages 53 and 54 of her dissertation that she did not intend to evaluate Read Naturally:

> *This study was not intended as an examination of the Read Naturally program. The lack of demonstrable standard score gains on the GORT Rate subtest for the Read Naturally group following this short term of training should not be construed as indicating that this instructional program is not effective in improving reading fluency. These children were not selected for their reading fluency problems – they were selected for their persisting writing problems.*

In addition, I cannot believe WWC does not have a length-of-study criteria. In the Chenault study ten sessions of twenty-five minutes (two weeks) is completely inadequate to evaluate a program. Students take two weeks to learn the steps and behavior to be successful in the Read Naturally strategy.

I strongly urge you not to post the Chenault study as a review of Read Naturally. Virginia Berninger and Belle Chenault both stated that their study was designed to evaluate Pay Attention and Writing.

I will send a more detailed analysis of WWC's use of Chenault as an evaluation of Read Naturally after I have had a chance to talk to the study's authors. I would appreciate the opportunity to discuss this issue with you. Please contact me at 651-286-8721.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

info@readnaturally.com

From:info@readnaturally.com
Sent:3 Aug 2010 13:03:34 -0500
To:What Works
Subject:Robert G St Pierre
Robert,

Now that I have talked to the authors of the Chenault study I will be able to prepare a more detailed analysis on why WWC should not post a review of Read Naturally based on the Chenault study.

Before the full analysis is complete, I've attached emails from the Chenault study authors clearly stating that the Chenault study should not be used as a review of Read Naturally.

Please begin an evaluation on the merits of removing the Chenault study. I will get you a full analysis by next week.

Thank you,

Thomas M Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Hi Tom,

After our conversation today, I did find lots of information I had saved about my dissertation process, but unfortunately I did not find specifics to answer your questions as to how many "read-alongs" students in the Read Naturally group did. As I said on the phone, I believe we used their initial reading pre-test results to inform placement in a Read Naturally level and I think we followed your recommended methods as to when they advanced through the books.

I did find a schedule and checked my calendar also to verify that each student had ten 25-minute training sessions with Read Naturally, two per week for 5 weeks. They then had the 55-minute writing instruction sessions as a large group with small group coaching, again two sessions per week for 5 weeks.

In the first phase of treatment, there were two students in each Read Naturally group, and groups varied as to whether both were on the same story and read together on initial "read alongs" or not. In my group, as I may have said, one of the students was quite low and so I did

choral readings with each student alone instead of as a group of three (two students and the instructor). I think that in the other groups, the students were close enough in skill to do "read-alongs" together, but I do not have specific data about this.

As we said on the phone today, I heartily agree with Dr. Berninger's statements that the purpose of our study was not to evaluate the effectiveness of Read Naturally, and with the statements on p. 53 and 54 of my dissertation to that effect: "The Read Naturally program is a widely recognized and widely used curriculum that has been appropriately validated in other studies. This study was not intended as an examination of the Read Naturally program." Using our study to evaluate Read Naturally or its effects is a misapplication of our data, in my opinion.

I enjoyed using the Read Naturally program in my dissertation research. The 4th through 6th graders in the study responded well to the interesting content of the stories. Since getting my PhD, I have often recommended Read Naturally to clients who are looking for reliable and engaging materials to build reading fluency. Thank you for making such a helpful program available!

I hope this information is useful to you and the Read Naturally program.

Best wishes, and thank you for contacting me about these matters.
Belle Chenault, PhD

Tom, I agree that you cannot use a contact control (which we knew would be good for the children and not a waste of their time but was not predicted to change the aspect of writing we were training) to draw conclusions about the effectiveness of the contact control. In fact there is an emerging literature on the fact that while reading and writing have some relationships there are aspects of writing not related to reading at all, which is why I think Joe Jenkins proposed using reading fluency training as a contact control.

I agree with Belle about how helpful Read Naturally is and we have used it for the purpose for which it was intended in other treatment studies.
In fact, I have donated my remaining copies to the Seattle Public Schools and they were thrilled to use the materials --since they were sitting lonely on our shelves because we do not currently have funding to do research.

Last week on the way home from the four week post op visit I stopped by the university and found a copy of Belle's published paper based on the dissertation research and asked the secretary to copy it and send it to you.
Please let me know when it arrives.

Also we are both ready to support you in the effort to help others see that they should not misinterpret the published research findings. That study was not about the merits of Read Naturally.

Ginger

info@readnaturally.com

From:info@readnaturally.com
Sent:3 Aug 2010 13:02:37 -0500
To:What Works
Subject:Mark Dynarski
Mark,

Now that I have talked to the authors of the Chenault study I will be able to prepare a more detailed analysis on why WWC should not post a review of Read Naturally based on the Chenault study.

Before the full analysis is complete, I've attached emails from the Chenault study authors clearly stating that the Chenault study should not be used as a review of Read Naturally.

Please begin an evaluation on the merits of removing the Chenault study. I will get you a full analysis by next week.

Thank you,

Thomas M Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Hi Tom,

After our conversation today, I did find lots of information I had saved about my dissertation process, but unfortunately I did not find specifics to answer your questions as to how many "read-alongs" students in the Read Naturally group did. As I said on the phone, I believe we used their initial reading pre-test results to inform placement in a Read Naturally level and I think we followed your recommended methods as to when they advanced through the books.

I did find a schedule and checked my calendar also to verify that each student had ten 25-minute training sessions with Read Naturally, two per week for 5 weeks. They then had the 55-minute writing instruction sessions as a large group with small group coaching, again two sessions per week for 5 weeks.

In the first phase of treatment, there were two students in each Read Naturally group, and groups varied as to whether both were on the same story and read together on initial "read alongs" or not. In my group, as I may have said, one of the students was quite low and so I did

choral readings with each student alone instead of as a group of three (two students and the instructor). I think that in the other groups, the students were close enough in skill to do "read-alongs" together, but I do not have specific data about this.

As we said on the phone today, I heartily agree with Dr. Berninger's statements that the purpose of our study was not to evaluate the effectiveness of Read Naturally, and with the statements on p. 53 and 54 of my dissertation to that effect: "The Read Naturally program is a widely recognized and widely used curriculum that has been appropriately validated in other studies. This study was not intended as an examination of the Read Naturally program." Using our study to evaluate Read Naturally or its effects is a misapplication of our data, in my opinion.

I enjoyed using the Read Naturally program in my dissertation research. The 4th through 6th graders in the study responded well to the interesting content of the stories. Since getting my PhD, I have often recommended Read Naturally to clients who are looking for reliable and engaging materials to build reading fluency. Thank you for making such a helpful program available!

I hope this information is useful to you and the Read Naturally program.

Best wishes, and thank you for contacting me about these matters.
Belle Chenault, PhD

Tom, I agree that you cannot use a contact control (which we knew would be good for the children and not a waste of their time but was not predicted to change the aspect of writing we were training) to draw conclusions about the effectiveness of the contact control. In fact there is an emerging literature on the fact that while reading and writing have some relationships there are aspects of writing not related to reading at all, which is why I think Joe Jenkins proposed using reading fluency training as a contact control.

I agree with Belle about how helpful Read Naturally is and we have used it for the purpose for which it was intended in other treatment studies.
In fact, I have donated my remaining copies to the Seattle Public Schools and they were thrilled to use the materials --since they were sitting lonely on our shelves because we do not currently have funding to do research.

Last week on the way home from the four week post op visit I stopped by the university and found a copy of Belle's published paper based on the dissertation research and asked the secretary to copy it and send it to you.
Please let me know when it arrives.

Also we are both ready to support you in the effort to help others see that they should not misinterpret the published research findings. That study was not about the merits of Read Naturally.

Ginger

info@readnaturally.com

From:info@readnaturally.com
Sent:11 Aug 2010 11:25:02 -0500
To:What Works
Subject:Robert G. St. Pierre
Robert G. St. Pierre,

I would appreciate it if you would read and respond to the following analysis on WWC using the Chenault study as an evaluation of Read Naturally.

The authors of the Chenault study are very clear that their study should not be used to evaluate Read Naturally. Belle Chenault stated in her dissertation:

> *This study was not intended as an examination of the Read Naturally program. The lack of demonstrable standard score gains on the GORT Rate subtest for the Read Naturally group following this short term of training should not be construed as indicating that this instructional program is not effective in improving reading fluency. These children were not selected for their reading fluency problems – they were selected for their persisting writing problems.*

In addition, Belle Chenault stated in a recent email:

> *"The Read Naturally program is a widely recognized and widely used curriculum that has been appropriately validated in other studies. This study was not intended as an examination of the Read Naturally program. Using our study to evaluate Read Naturally or its effects is a misapplication of our data."*

Belle Chenault's advisor, Dr. Virginia Berninger, also stated that What Works Clearinghouse (WWC) using the Chenault study as an evaluation of Read Naturally was a misinterpretation of the Chenault study's published research findings. Dr. Berninger stated:

> *You cannot use a contact control (which we knew would be good for the children and not a waste of their time but was not predicted to change the aspect of writing we were training) to draw conclusions about the effectiveness of the contact control. In fact there is an emerging literature on the fact that, while reading and writing have some relationships, there are aspects of writing not related to reading at all, which is why I think Joe Jenkins proposed using reading fluency training [Read Naturally] as a contact control.*

According to Dr. Berninger, "In no way can this kind of experimental design, which is well established in educational science, be considered a randomized controlled evaluation of the contact control for its treatment effectiveness."

The Chenault study used Read Naturally as a contact control. Consequently, this study cannot be used to draw any conclusions about Read Naturally.

Also, there were significant implementation discrepancies between Read Naturally's recommended process and the way the Chenault study used Read Naturally. The Chenault study did not follow Read Naturally's process in the following ways:

1. Minimal Time Spent Reading. Read Naturally recommends that the minimum standard for students to engage in the Read Naturally strategy is at least three times a

week. In the Chenault study, students met two times a week for 25 minutes a session, for just five weeks.

2. Length of Study Too Short. When introducing students to the Read Naturally strategy, Read Naturally expects students to need six to ten sessions to learn the steps of the strategy in order to make progress. In the Chenault study, students only met for ten, 25-minute sessions total (only two weeks if the students had met each day of the week). This was enough time to learn the steps but not nearly enough time to demonstrate progress.
3. Placement Process Not Followed. Read Naturally recommends a specific placement procedure based on words correct per minute (WCPM) on one minute timed readings of passages from the Read Naturally placement packet. In the Chenault study, students were placed based on the pretest assessment measures. These measures would have been a good starting point for using the Read Naturally placement process but would not have been sufficient. Using the passages in the placement packet allows teachers to find a precise level at which to start the student, so the student will be challenged but not frustrated. Typically, when starting at the initial placement level, the student should need three read alongs and three to six one-minute practice reads before they could meet the four criteria to pass the story.
4. Study Fails to Individualize. Read Naturally recommends that students work individually in the strategy to maximize their progress in reading and to develop independent skills. In the Chenault study, most of the students did a choral reading with another student and a teacher for the read along portion of the strategy. Typically, this method would hold back the progress of at least one of the students. Also, most students benefit from the gradually increasing speeds on Read Naturally's audio recordings, which were not used in the study.
5. Comprehension Pass Criteria Skipped. Answering all the comprehension questions correctly is one of four criteria to pass in the Read Naturally strategy. If a student misses a question, the student has to go back and find the answer in the text. In the Chenault study, only three of the passing criteria were required (meet goal rate, make three or fewer errors, and read with good expression).

What Works Clearinghouse (WWC) should follow the recommendation of the Chenault study authors and remove the study as an evaluation of Read Naturally from the WWC website. Please contact me to discuss this issue. (651-286-8721)

Sincerely,

Thomas M. Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Maximize the benefits of your Read Naturally program. Attend one of our full-day training seminars near you!

Join Read Naturally on Facebook

Find us on
Facebook

info@readnaturally.com

From:info@readnaturally.com
Sent:11 Aug 2010 11:26:05 -0500
To:What Works
Subject:Mark Dynarski
Mark Dynarski,

I would appreciate it if you would read and respond to the following analysis on WWC using the Chenault study as an evaluation of Read Naturally.

The authors of the Chenault study are very clear that their study should not be used to evaluate Read Naturally. Belle Chenault stated in her dissertation:

> *This study was not intended as an examination of the Read Naturally program. The lack of demonstrable standard score gains on the GORT Rate subtest for the Read Naturally group following this short term of training should not be construed as indicating that this instructional program is not effective in improving reading fluency. These children were not selected for their reading fluency problems – they were selected for their persisting writing problems.*

In addition, Belle Chenault stated in a recent email:

> *"The Read Naturally program is a widely recognized and widely used curriculum that has been appropriately validated in other studies. This study was not intended as an examination of the Read Naturally program. Using our study to evaluate Read Naturally or its effects is a misapplication of our data."*

Belle Chenault's advisor, Dr. Virginia Berninger, also stated that What Works Clearinghouse (WWC) using the Chenault study as an evaluation of Read Naturally was a misinterpretation of the Chenault study's published research findings. Dr. Berninger stated:

> *You cannot use a contact control (which we knew would be good for the children and not a waste of their time but was not predicted to change the aspect of writing we were training) to draw conclusions about the effectiveness of the contact control. In fact there is an emerging literature on the fact that, while reading and writing have some relationships, there are aspects of writing not related to reading at all, which is why I think Joe Jenkins proposed using reading fluency training [Read Naturally] as a contact control.*

According to Dr. Berninger, "In no way can this kind of experimental design, which is well established in educational science, be considered a randomized controlled evaluation of the contact control for its treatment effectiveness."

The Chenault study used Read Naturally as a contact control. Consequently, this study cannot be used to draw any conclusions about Read Naturally

Also, there were significant implementation discrepancies between Read Naturally's recommended process and the way the Chenault study used Read Naturally. The Chenault study did not follow Read Naturally's process in the following ways:

1. Minimal Time Spent Reading. Read Naturally recommends that the minimum standard for students to engage in the Read Naturally strategy is at least three times a

week. In the Chenault study, students met two times a week for 25 minutes a session, for just five weeks.

2. Length of Study Too Short. When introducing students to the Read Naturally strategy, Read Naturally expects students to need six to ten sessions to learn the steps of the strategy in order to make progress. In the Chenault study, students only met for ten, 25-minute sessions total (only two weeks if the students had met each day of the week). This was enough time to learn the steps but not nearly enough time to demonstrate progress.
3. Placement Process Not Followed. Read Naturally recommends a specific placement procedure based on words correct per minute (WCPM) on one minute timed readings of passages from the Read Naturally placement packet. In the Chenault study, students were placed based on the pretest assessment measures. These measures would have been a good starting point for using the Read Naturally placement process but would not have been sufficient. Using the passages in the placement packet allows teachers to find a precise level at which to start the student, so the student will be challenged but not frustrated. Typically, when starting at the initial placement level, the student should need three read alongs and three to six one-minute practice reads before they could meet the four criteria to pass the story.
4. Study Fails to Individualize. Read Naturally recommends that students work individually in the strategy to maximize their progress in reading and to develop independent skills. In the Chenault study, most of the students did a choral reading with another student and a teacher for the read along portion of the strategy. Typically, this method would hold back the progress of at least one of the students. Also, most students benefit from the gradually increasing speeds on Read Naturally's audio recordings, which were not used in the study.
5. Comprehension Pass Criteria Skipped. Answering all the comprehension questions correctly is one of four criteria to pass in the Read Naturally strategy. If a student misses a question, the student has to go back and find the answer in the text. In the Chenault study, only three of the passing criteria were required (meet goal rate, make three or fewer errors, and read with good expression).

What Works Clearinghouse (WWC) should follow the recommendation of the Chenault study authors and remove the study as an evaluation of Read Naturally from the WWC website. Please contact me to discuss this issue. (651-286-8721)

Sincerely,

Thomas M. Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Maximize the benefits of your Read Naturally program. Attend one of our full-day training seminars near you!

Join Read Naturally on Facebook

Find us on
Facebook

February 26, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013

Dear Mr. Ihnot:

Thank you for your inquiries concerning the What Works Clearinghouse (WWC) reviews of four studies of Read Naturally®—Kemp (2006), Denton et al. (2004), Hancock (2002), and Chenault et al. (2006)—and for your email regarding my response to Dr. Kevin Feldman's inquiry about the WWC reviews of studies of Read Naturally®. In response to your concerns, we conducted an independent quality review. The WWC quality review team responds to concerns raised by study authors, curriculum developers, or other relevant parties about WWC reviews published on our website. When a quality review is conducted, a researcher who was not involved in the initial reviews undertakes an independent assessment of the studies in question. The researcher also investigates the procedures used and decisions made during the original reviews of the studies. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the WWC.

The quality review considered the concerns from your emails regarding study inclusion, implementation, student sample, and outcomes.

Study Inclusion

The WWC followed protocol in choosing to review these four studies. The WWC screens studies based on a number of factors, including relevancy and methodology criteria (as described in the WWC Procedures and Standards Handbook, pages 8–10, available at http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19). The WWC does not screen based on whether the author(s) explicitly intended the study as an evaluation of an intervention, but rather whether the study presents a primary analysis of the effect of an intervention. This screening process allows for the review of a wide range of relevant and methodologically sound studies.

Thank you for bringing to my attention the studies by Mesa (2004) and Tucker (2010). The Tucker (2010) study will be included in our literature search for updating the Intervention Report

for Read Naturally®, reviewed under the Beginning Reading Evidence Review Protocol. In addition, as you note, the Mesa (2004) study was included in the original review of Read Naturally®, but will be reviewed again in an updated review.

Implementation

The WWC followed protocol in the manner in which the four studies are described in WWC publications. The WWC Handbook says, "The WWC makes no adjustments or corrections for variations in implementation of the intervention; however, if a study meets standards and is included in an intervention report, descriptions of implementation are provided in the report appendices to provide context for the findings" (page 16). This approach is appropriate because there is no standard metric for fidelity to intervention design. Thus, the WWC includes studies with variation in fidelity and does not evaluate implementation fidelity. The quality review team verified that variations in implementation that are noted in the four studies and that may affect the interpretation of findings were properly included in the WWC publications.

Regarding the Kemp (2006) study, your emails raised a concern about implementation of the Read Naturally® program. The quality review noted that the study does not provide any indication of deviation from program design. Specifically, the study states, "… it could be concluded that all teachers implemented the Read Naturally® program as prescribed" (page 40). Based on this information, the quality review concluded that there was no evidence that deviations in implementation should have been noted in the WWC Intervention Report. Regarding your specific concern about the teacher-to-student ratio, the study does not note any deviation from the Read Naturally® program in this area and does not provide any information on the ratio.

You also expressed concern over deviations in the implementation of Read Naturally® in the Denton et al. (2004) study. The quality review found that the inclusion of additional activities was noted in the study and reported in the WWC Intervention Report (see footnote 7 on page 3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). The other potential deviations that you noted (accountability to pass criteria, use of audio tapes, and the number of sessions per week) were not identified in the study and are therefore not noted in the WWC publication.

Regarding your concerns that the Read Naturally® strategy was not implemented fully for the Hancock (2002) study, the quality review found that WWC procedures were followed in the WWC Intervention Report. The WWC contacted Hancock regarding the implementation of Read Naturally® in the 2002 study. Her response indicated that the study excluded Read Naturally®'s pre-reading vocabulary instruction component and the placement system to individualize instruction. The WWC Intervention Report was revised to note these deviations in implementation (see footnote 4 on page 2, available at http://ies.ed.gov/ncee/wwc/InterventionReport.aspx?sid=407). Other potential concerns with implementation described in your email were not reported by the author and were not included in the revised WWC Intervention Report.

The quality review also investigated concerns that you raised in prior emails. The email from June 2012 regarding Dr. Feldman's questions was addressed in my letter to Dr. Feldman; I will not

repeat the quality review findings here. Your email from May 2012 included some text that was reported to be a quote from Hancock's email to the WWC, but the quoted statements were not part of Hancock's email text. The misreported quote is also found on the Read Naturally® website at http://www.readnaturally.com/approach/hancock.htm.

Regarding the Chenault et al. (2006) study, you expressed concern with the implementation of Read Naturally® related to the selection of students and the limited number of sessions. The quality review found that the WWC Intervention Report appropriately notes both of these factors (see pages 2–3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=409).

Student Sample

You also raised a concern that the Kemp (2006) study included a large share of students who were above the 50th percentile in reading fluency who were not "at-risk" readers. The quality review verified that the WWC Intervention Report accurately states that the study population for the results reported by the WWC was third-grade English language learner (ELL) students (page 2, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). Furthermore, the quality review noted that Read Naturally® program materials state that the program "… products help students of all ages and abilities—whether special education, ELL, Title I, or mainstream—to improve their reading skills, gain confidence, and transform their attitudes toward school and learning" (available at http://www.readnaturally.com/products/default.htm).

Finally, regarding the differences in pre-test performance among native-speaking students in the treatment and control groups, the quality review found that these concerns are not relevant to this WWC publication, which reports on findings only for English language learner students. Furthermore, the quality review concluded that the WWC appropriately restricted the reported findings to impacts assessed relative to the comparison group, and thus did not include achievement gains in oral reading fluency.

Outcomes

Regarding the Denton et al. (2004) study, you expressed concern that the study did not assess fluency. The quality review noted that the WWC Intervention Report clearly states that the study assessed reading achievement (page 4, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408), an outcome of relevance to the WWC ELL topic area. The quality review noted that the Read Naturally® program materials state there are several components of the program beyond fluency that are related to reading achievement: "Our industry-leading programs develop and support the five essential components of reading, identified by the National Reading Panel: phonemic awareness, phonics, fluency, vocabulary, and comprehension" (see "The Read Naturally Strategy," available at http://www.readnaturally.com/approach/default.htm).

I hope that this letter has addressed your concerns. You requested that we meet to discuss the WWC reviews of Read Naturally®. However, WWC policy requires all correspondence in writing. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov.

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

cc: Joy Lesnick, IES

Sent:26 Feb 2013 20:09:42 +0000
To:'kramy@readnaturally.com';'info@readnaturally.com'
Subject:What Works Clearinghouse (QR2010013)
Attachments:QRT2010013.pdf
Dear Mr. Ihnot:

Attached is a response to the questions you raised in your inquiries concerning the What Works Clearinghouse reviews of studies of Read Naturally®.

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/.

March 5, 2013

Dear Jill Constantine,

Thank you for your letter dated 2/26/13. I appreciate that you took the time to address the concerns I have raised and to conduct some research about Read Naturally, Inc.

I am responding in hopes that we can move forward with a plan that stays true to WWC's mission to provide educators with credible and reliable information. I share this goal, which is why it's important for me to clarify some misunderstandings.

Read Naturally, Inc. is a company that offers various products targeting many aspects of reading instruction. The statements from our website that you quoted in your letter refer to Read Naturally as a company that offers many different products targeting students of various abilities and with various needs. Read Naturally® is also the name of a specific fluency-building strategy, which I believe has been a source of confusion for researchers, educators, and WWC.

When an educator googles "Read Naturally," the second search result is WWC's statement: "**Read Naturally** was found to have no discernible effects on fluency and reading comprehension."[1] The four studies used to support this statement did not evaluate Read Naturally, Inc., the company as a whole, nor did the studies evaluate the Read Naturally fluency-building strategy, as all four authors admit that they did not use it as prescribed. Thus, the "Read Naturally" in WWC's statement cannot accurately refer to the company or the strategy. It also can't refer to a specific product, as we do not offer a product called "Read Naturally." Our fluency-building products are called Encore, Software Edition, and Read Live.

When researchers call a method or group "Read Naturally" but do not use our complete fluency-building strategy as prescribed, they are not using trademarked Read Naturally name correctly. This has resulted in WWC misinterpreting the study findings and publishing a damaging statement about Read Naturally, Inc. I have spent considerable time and expense over the past seven years trying to clarify this confusion for educators.

As an immediate solution, I am requesting that WWC remove or change the statement "Read Naturally was found to have no discernible effects on fluency and reading comprehension." First, the statement needs to clarify the nebulous use of "Read Naturally." In fairness, the "Read Naturally" in this statement is referring to "leveled passages published by Read Naturally, Inc." and should be replaced as such. Second, please reconsider the phrase "no discernible effects," as this does not seem a fitting description of the study findings. Implementation issues aside, the WWC improvement index for the Hancock study shows an average of +8 percentile points for fluency. The intervention group had a 1.68 WCPM weekly gain, which was greater than the control group and greater than the average weekly gain for second graders based on oral reading fluency norms. In the Kemp study, the so-called Read Naturally group also had greater weekly fluency gains than the SSR group, and the gains were above the average for third graders.

---

[1] Read Naturally search result #2:
Read Naturally®: What Works Clearinghouse
ies.ed.gov › Home › Topics › Literacy
Read Naturally was found to have no discernible effects on fluency and reading comprehension.

In summary, I request one or some combination of the following options to be immediately implemented:

1. Remove any statement referring to Read Naturally completely.
2. Replace "Read Naturally" with "leveled passages published by Read Naturally, Inc." and evaluate the current data on the WWC website for the Hancock and Kemp studies to formulate new summary statement to replace "no discernible effects."

In addition to changing this statement, I would appreciate it if you would contact the authors of all four studies in question so that they may verify for you why their studies should not be reviewed as evaluations of Read Naturally, Inc., the Read Naturally strategy, or a specific Read Naturally product. See Addendum 1 for details about these studies, including specific responses to statements in your letter to me on 2/26/13. This addendum includes the authors' contact information.

Finally, WWC has promised me a reevaluation of its reviews. This reevaluation should take into account the statements the study authors have made to Read Naturally, Inc., as well as the statements they make to you when you contact them. Additionally, the reevaluation should take into account the fact that another United States Department of Education entity, the National Center on Intensive Intervention (NCII), has reviewed three studies that implemented the Read Naturally strategy correctly. These studies show that the Read Naturally strategy has a significant impact on reading achievement in fluency, accuracy, rate and comprehension. See Addendum 2 for more details about these studies and other well-designed case studies that demonstrate the effectiveness of the Read Naturally strategy.

I appreciate your continued correspondence with me as we work to resolve this issue in a way that is fair to Read Naturally, Inc., WWC, and the educators we both serve. I will await your response letting me know when you will change WWC's published statement. I would also appreciate a timeframe for when I can expect WWC to reevaluate its review of our strategy and/or specific products. I look forward to a fair and open-minded treatment of this issue and hope to continue this discussion as soon as possible.

Sincerely,

Tom Ihnot
CEO

What Works

From:What Works
Sent:5 Mar 2013 15:56:18 -0500
To:WhatWorks
Cc:Sakari Morvey
Subject:FW: Read Naturally - Tom Ihnot (WWCPC 3661)
Attachments:2ndConstantine.march.2013.pdf, Addendum 1.pdf, Addendum 2.pdf

Hello Kate and Sakari,

We received another email from Tom Ihnot. I have added it to Issue 3661. Sakari, I was unsure if it should be added to a topic area library as well, since it's addressed to the help desk. Let me know if you would like me to handle the issue differently.

Thanks,

Ariel Duran
(202) 250-3506

From: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Tuesday, March 05, 2013 3:33 PM
To: What Works
Subject: Read Naturally - Tom Ihnot

Jill,

Please review the three attachments that are a response to your February 26, 2013 letter.

Thank you,
Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: Anne Hauth (b)(6)
Sent: Tuesday, March 05, 2013 1:44 PM
To: info@readnaturally.com
Subject: for Tom

Hi Tom,

I updated the letter to Jill to include the footnote about the google search result, and I updated addendum 1 to include the screen shot of the Hancock email. I'm attaching final versions of all three pieces: letter to Jill, addendum 1, and addendum 2. I'm emailing you the word documents as well as PDFs. You can send Jill either the word docs or the PDFs depending on what you'd prefer. Sometimes PDFs are easier since not everyone has the same version of Word.

Let me know if you need anything else or have questions.

Thanks!
Anne

Addendum 1

The passages published by Read Naturally, Inc. are an attractive tool for researchers because they are carefully leveled short passages at a wide range of reading levels. The Hancock, Kemp, Denton, and Chenault studies all used passages published by Read Naturally, Inc., but none used the Read Naturally® strategy. Therefore, they cannot be used to evaluate the Read Naturally strategy or a specific Read Naturally, Inc. product.

Hancock Study

In addressing the Hancock study, WWC states in the footnotes that Hancock indicated her study excluded the Read Naturally strategy's pre-reading vocabulary instruction and the placement system to individualize instruction. These categories represent a wide range of major deviations, and there are more that should be included:

- Prediction step
- Adjusting goals and levels
- Accountability for passing

However, merely mentioning the Hancock study deviations in footnotes will not help inform educators. They do not read the footnotes. The vast majority will only see "Read Naturally was found to have no discernible effects on fluency and reading comprehension" when they "google" Read Naturally.

Carrie Hancock states in her dissertation that critical aspects of the Read Naturally strategy were not part of her study and that all aspects of the Read Naturally strategy should be included in fluency interventions. Since Hancock did not implement the Read Naturally strategy, her study cannot be used to evaluate the Read Naturally strategy as a primary analysis of the effect of an intervention.

This is a screenshot of the email Carrie Hancock sent to Read Naturally, Inc., which verifies that her study did not evaluate the Read Naturally strategy:

From: (b)(6)
Sent: Monday, July 16, 2007 10:15 PM
To: info@readnaturally.com
Subject: dissertation clarification

Tom,

Per our phone conversation this morning, I wanted to calrify a couple of issues regarding my 2002 dissertation titled, "Acclerating Reading Trajectories: The Effects of Dynamic Research-Based Instruction". While I used Read Naturally materials, I did not fully implement the Read Naturally strategy and my study was not intended to evaluate the Read Naturally strategy. Rather, the purpose was to determine the impact of ongoing supplemental fluency practice on 2nd grade students' rates of learning to read.

Thank you,

Carrie Hancock Ph.D.

Carrie Hancock can be contacted at (b)(6) to verify the deviations of her study and the fact that she only used passages published by Read Naturally, Inc. and not the Read Naturally strategy.

Kemp Study

Concerning the Kemp study, WWC quotes the study's statement: "It could be concluded that all teachers implemented the Read Naturally program as prescribed." This statement means that Kemp used passages published by Read Naturally, Inc. in the model prescribed in her study. However, her model was not the Read Naturally strategy. Consequently, she will not list the deviations from the Read Naturally strategy since her study is not implementing the Read Naturally strategy.

On page 74 of her dissertation, Kemp clearly states that her study was not appropriate for implementing the Read Naturally strategy since the vast majority of the students were not at risk. Of course, she did not implement the Read Naturally strategy and consequently her study should not be used as a primary analysis of an intervention.

In spite of the Kemp study's design flaws, the group using passages from Read Naturally, Inc. did make significant gains in fluency. These students averaged a gain of 1.5 WCPM (words correct per minute) each week. Oral reading fluency norms show that the average fluency gain per week for third grade students is 1.1 WCPM. This data indicates that the students using the passages from Read Naturally, Inc. were accelerating their progress.

Susan Kemp can be contacted at (b)(6) or (b)(6), and more information can be found here: http://www.readnaturally.com/approach/kemp.htm

Denton Study

Concerning your comments on the Denton study, you cite the statement on our website: "Our industry-leading programs develop and support the five essential components …" The Read Naturally strategy and its flagship programs, Masters Edition, Encore, Software Edition, and Read Live develop fluency, support vocabulary, and promote comprehension. We offer other programs that develop phonics and support phonemic awareness: Group and Tutoring Edition (GATE), Word Warm-ups, and Signs for Sounds. The Denton study did not come close to using the Read Naturally strategy and did not use GATE, Word Warm-ups or Signs for Sounds for phonics or phonemic awareness.

You can contact Carolyn Denton at carolyn.a.denton@uth.tmc.edu. You can also contact the Denton study advisor, Dr. Jan Hasbrouck at (b)(6) or (b)(6). More information can be found here: http://www.readnaturally.com/approach/denton.htm

Chenault Study

The Chenault study authors have strongly stated their objections for using their study as an evaluation of the Read Naturally strategy. You can contact Belle Chenault at (b)(6) or (b)(6) and study advisor Dr. Virginia Berninger at vwb@u.washington.edu or 206-616-6372. More information can be found here: http://www.readnaturally.com/approach/chenault.htm

Addendum 2

The National Center on Intensive Intervention (NCII), has reviewed three studies that intended to evaluate the Read Naturally® strategy and implemented it properly. These studies show that the Read Naturally strategy has a significant impact on reading achievement in fluency, accuracy, rate and comprehension. NCII reviewed the Heistad study, the Christ study, and the Tucker study.

1. The Heistad study shows moderate effect size impact on reading comprehension. Because the study lasted a full school year, it was long enough for comprehension assessment tools to pick up students' comprehension gains. Complete results of the Heistad study can be found here: http://www.readnaturally.com/pdf/heistad-study_4schools.pdf.

2. The Christ study was a ten-week study that shows a significant effect size impact on fluency and accuracy. Dr. Christ administered a comprehension assessment but did not expect to see an impact because good comprehension assessment tools are not sensitive enough to pick up gains in only ten weeks. Complete results of the Christ study can be found here: http://www.readnaturally.com/pdf/UofMnReadNaturallyStudy.pdf.

3. The Tucker study shows moderate to large effect size impact in rate, accuracy and fluency. More information on the Tucker study can be found here: http://www.readnaturally.com/company/news_ncii.htm.

Additionally, the following six control group studies demonstrate the Read Naturally strategy's effectiveness. Please review this compelling evidence.

Case 1: Original Study, Minneapolis, MN
http://www.readnaturally.com/approach/case1.htm

Case 5: "Throwing Sixth Graders a Lifeline"
http://www.readnaturally.com/approach/case5.htm

Case 6: Special Education Students, Huron County, MI
http://www.readnaturally.com/approach/case6.htm

Case 7: First Graders, South Forsyth County, GA
http://www.readnaturally.com/approach/case7.htm

Case 10: Third Graders, Southern California
http://www.readnaturally.com/approach/case10.htm

Case 11: Second Graders, Elk River, MN
http://www.readnaturally.com/approach/case11.htm

Susan Kemp

From:Susan Kemp
Sent:12 Mar 2013 09:31:30 -0700
To:What Works
Subject:RE: Read Naturally - Tom Ihnot

Dear Ms. Constantine

Please see the email below. I in no way wish to be involved in this situation. My study was looking at repeated reading vs sustained silent reading. I used Read Naturally as the tool to investigate the effects of repeated reading.

Thank you

Dr. Susan Kemp

From: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Wednesday, March 06, 2013 8:32 AM
To: Susan Kemp
Subject: Read Naturally - Tom Ihnot

Dr Susan Kemp,

I am still communicating with Jill Constantine of What Works Clearinghouse (WWC) to persuade them to change their posted summary statement about Read Naturally. I am hopeful there will be a change in the statement because the Office of Representative John Kline, chairman of the U.S. House of Representatives Education Committee has taken an interest in WWC policies.

I have attached my most recent letter and two addendums to Jill Constantine to provide you the latest background information. I would appreciate it if you could email Jill Constantine (whatworks@mathematica-mpr.com) to let her know that your study should not be used as an evaluation of the Read Naturally strategy. The statement is inaccurate, misinforms educators and unfairly damages Read Naturally.

Thank you,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally®
CELEBRATING
20 YEARS

March 5, 2013

Dear Jill Constantine,

Thank you for your letter dated 2/26/13. I appreciate that you took the time to address the concerns I have raised and to conduct some research about Read Naturally, Inc.

I am responding in hopes that we can move forward with a plan that stays true to WWC's mission to provide educators with credible and reliable information. I share this goal, which is why it's important for me to clarify some misunderstandings.

Read Naturally, Inc. is a company that offers various products targeting many aspects of reading instruction. The statements from our website that you quoted in your letter refer to Read Naturally as a company that offers many different products targeting students of various abilities and with various needs. Read Naturally® is also the name of a specific fluency-building strategy, which I believe has been a source of confusion for researchers, educators, and WWC.

When an educator googles "Read Naturally," the second search result is WWC's statement: "**Read Naturally** was found to have no discernible effects on fluency and reading comprehension."[1] The four studies used to support this statement did not evaluate Read Naturally, Inc., the company as a whole, nor did the studies evaluate the Read Naturally fluency-building strategy, as all four authors admit that they did not use it as prescribed. Thus, the "Read Naturally" in WWC's statement cannot accurately refer to the company or the strategy. It also can't refer to a specific product, as we do not offer a product called "Read Naturally." Our fluency-building products are called Encore, Software Edition, and Read Live.

When researchers call a method or group "Read Naturally" but do not use our complete fluency-building strategy as prescribed, they are not using trademarked Read Naturally name correctly. This has resulted in WWC misinterpreting the study findings and publishing a damaging statement about Read Naturally, Inc. I have spent considerable time and expense over the past seven years trying to clarify this confusion for educators.

As an immediate solution, I am requesting that WWC remove or change the statement "Read Naturally was found to have no discernible effects on fluency and reading comprehension." First, the statement needs to clarify the nebulous use of "Read Naturally." In fairness, the "Read Naturally" in this statement is referring to "leveled passages published by Read Naturally, Inc." and should be replaced as such. Second, please reconsider the phrase "no discernible effects," as this does not seem a fitting description of the study findings. Implementation issues aside, the WWC improvement index for the Hancock study shows an average of +8 percentile points for fluency. The intervention group had a 1.68 WCPM weekly gain, which was greater than the control group and greater than the average weekly gain for second graders based on oral reading fluency norms. In the Kemp study, the so-called Read Naturally group also had greater weekly fluency gains than the SSR group, and the gains were above the average for third graders.

---

[1] Read Naturally search result #2:
Read Naturally®: What Works Clearinghouse
ies.ed.gov › Home › Topics › Literacy
Read Naturally was found to have no discernible effects on fluency and reading comprehension.

In summary, I request one or some combination of the following options to be immediately implemented:

1. Remove any statement referring to Read Naturally completely.
2. Replace "Read Naturally" with "leveled passages published by Read Naturally, Inc." and evaluate the current data on the WWC website for the Hancock and Kemp studies to formulate new summary statement to replace "no discernible effects."

In addition to changing this statement, I would appreciate it if you would contact the authors of all four studies in question so that they may verify for you why their studies should not be reviewed as evaluations of Read Naturally, Inc., the Read Naturally strategy, or a specific Read Naturally product. See Addendum 1 for details about these studies, including specific responses to statements in your letter to me on 2/26/13. This addendum includes the authors' contact information.

Finally, WWC has promised me a reevaluation of its reviews. This reevaluation should take into account the statements the study authors have made to Read Naturally, Inc., as well as the statements they make to you when you contact them. Additionally, the reevaluation should take into account the fact that another United States Department of Education entity, the National Center on Intensive Intervention (NCII), has reviewed three studies that implemented the Read Naturally strategy correctly. These studies show that the Read Naturally strategy has a significant impact on reading achievement in fluency, accuracy, rate and comprehension. See Addendum 2 for more details about these studies and other well-designed case studies that demonstrate the effectiveness of the Read Naturally strategy.

I appreciate your continued correspondence with me as we work to resolve this issue in a way that is fair to Read Naturally, Inc., WWC, and the educators we both serve. I will await your response letting me know when you will change WWC's published statement. I would also appreciate a timeframe for when I can expect WWC to reevaluate its review of our strategy and/or specific products. I look forward to a fair and open-minded treatment of this issue and hope to continue this discussion as soon as possible.

Sincerely,

Tom Ihnot
CEO

Addendum 1

The passages published by Read Naturally, Inc. are an attractive tool for researchers because they are carefully leveled short passages at a wide range of reading levels. The Hancock, Kemp, Denton, and Chenault studies all used passages published by Read Naturally, Inc., but none used the Read Naturally® strategy. Therefore, they cannot be used to evaluate the Read Naturally strategy or a specific Read Naturally, Inc. product.

Hancock Study

In addressing the Hancock study, WWC states in the footnotes that Hancock indicated her study excluded the Read Naturally strategy's pre-reading vocabulary instruction and the placement system to individualize instruction. These categories represent a wide range of major deviations, and there are more that should be included:

- Prediction step
- Adjusting goals and levels
- Accountability for passing

However, merely mentioning the Hancock study deviations in footnotes will not help inform educators. They do not read the footnotes. The vast majority will only see "Read Naturally was found to have no discernible effects on fluency and reading comprehension" when they "google" Read Naturally.

Carrie Hancock states in her dissertation that critical aspects of the Read Naturally strategy were not part of her study and that all aspects of the Read Naturally strategy should be included in fluency interventions. Since Hancock did not implement the Read Naturally strategy, her study cannot be used to evaluate the Read Naturally strategy as a primary analysis of the effect of an intervention.

This is a screenshot of the email Carrie Hancock sent to Read Naturally, Inc., which verifies that her study did not evaluate the Read Naturally strategy:

From: (b)(6)
Sent: Monday, July 16, 2007 10:15 PM
To: info@readnaturally.com
Subject: dissertation clarification

Tom,

Per our phone conversation this morning, I wanted to calrify a couple of issues regarding my 2002 dissertation titled, "Acclerating Reading Trajectories: The Effects of Dynamic Research-Based Instruction". While I used Read Naturally materials, I did not fully implement the Read Naturally strategy and my study was not intended to evaluate the Read Naturally strategy. Rather, the purpose was to determine the impact of ongoing supplemental fluency practice on 2nd grade students' rates of learning to read.

Thank you,

Carrie Hancock Ph.D.

Carrie Hancock can be contacted at (b)(6) to verify the deviations of her study and the fact that she only used passages published by Read Naturally, Inc. and not the Read Naturally strategy.

Kemp Study

Concerning the Kemp study, WWC quotes the study's statement: "It could be concluded that all teachers implemented the Read Naturally program as prescribed." This statement means that Kemp used passages published by Read Naturally, Inc. in the model prescribed in her study. However, her model was not the Read Naturally strategy. Consequently, she will not list the deviations from the Read Naturally strategy since her study is not implementing the Read Naturally strategy.

On page 74 of her dissertation, Kemp clearly states that her study was not appropriate for implementing the Read Naturally strategy since the vast majority of the students were not at risk. Of course, she did not implement the Read Naturally strategy and consequently her study should not be used as a primary analysis of an intervention.

In spite of the Kemp study's design flaws, the group using passages from Read Naturally, Inc. did make significant gains in fluency. These students averaged a gain of 1.5 WCPM (words correct per minute) each week. Oral reading fluency norms show that the average fluency gain per week for third grade students is 1.1 WCPM. This data indicates that the students using the passages from Read Naturally, Inc. were accelerating their progress.

Susan Kemp can be contacted at (b)(6) or (b)(6), and more information can be found here: http://www.readnaturally.com/approach/kemp.htm

Denton Study

Concerning your comments on the Denton study, you cite the statement on our website: "Our industry-leading programs develop and support the five essential components …" The Read Naturally strategy and its flagship programs, Masters Edition, Encore, Software Edition, and Read Live develop fluency, support vocabulary, and promote comprehension. We offer other programs that develop phonics and support phonemic awareness: Group and Tutoring Edition (GATE), Word Warm-ups, and Signs for Sounds. The Denton study did not come close to using the Read Naturally strategy and did not use GATE, Word Warm-ups or Signs for Sounds for phonics or phonemic awareness.

You can contact Carolyn Denton at carolyn.a.denton@uth.tmc.edu. You can also contact the Denton study advisor, Dr. Jan Hasbrouck at (b)(6) or (b)(6). More information can be found here: http://www.readnaturally.com/approach/denton.htm

Chenault Study

The Chenault study authors have strongly stated their objections for using their study as an evaluation of the Read Naturally strategy. You can contact Belle Chenault at (b)(6) or (b)(6) and study advisor Dr. Virginia Berninger at vwb@u.washington.edu or 206-616-6372. More information can be found here: http://www.readnaturally.com/approach/chenault.htm

Addendum 2

The National Center on Intensive Intervention (NCII), has reviewed three studies that intended to evaluate the Read Naturally® strategy and implemented it properly. These studies show that the Read Naturally strategy has a significant impact on reading achievement in fluency, accuracy, rate and comprehension. NCII reviewed the Heistad study, the Christ study, and the Tucker study.

1. The Heistad study shows moderate effect size impact on reading comprehension. Because the study lasted a full school year, it was long enough for comprehension assessment tools to pick up students' comprehension gains. Complete results of the Heistad study can be found here: http://www.readnaturally.com/pdf/heistad-study_4schools.pdf.

2. The Christ study was a ten-week study that shows a significant effect size impact on fluency and accuracy. Dr. Christ administered a comprehension assessment but did not expect to see an impact because good comprehension assessment tools are not sensitive enough to pick up gains in only ten weeks. Complete results of the Christ study can be found here: http://www.readnaturally.com/pdf/UofMnReadNaturallyStudy.pdf.

3. The Tucker study shows moderate to large effect size impact in rate, accuracy and fluency. More information on the Tucker study can be found here: http://www.readnaturally.com/company/news_ncii.htm.

Additionally, the following six control group studies demonstrate the Read Naturally strategy's effectiveness. Please review this compelling evidence.

Case 1: Original Study, Minneapolis, MN
http://www.readnaturally.com/approach/case1.htm

Case 5: "Throwing Sixth Graders a Lifeline"
http://www.readnaturally.com/approach/case5.htm

Case 6: Special Education Students, Huron County, MI
http://www.readnaturally.com/approach/case6.htm

Case 7: First Graders, South Forsyth County, GA
http://www.readnaturally.com/approach/case7.htm

Case 10: Third Graders, Southern California
http://www.readnaturally.com/approach/case10.htm

Case 11: Second Graders, Elk River, MN
http://www.readnaturally.com/approach/case11.htm

info@readnaturally.com

From:info@readnaturally.com
Sent:20 Mar 2013 20:05:16 +0000
To:What Works
Subject:Read Naturally - Tom Ihnot
Attachments:2ndConstantine.march.2013.pdf, Addendum 1.pdf, Addendum 2.pdf

Dear Jill Constantine,

Thank you for the courtesy copy of the Adolescent Literacy review of the Heistad study.

Will this result in a change to the following general description that appears when you google Read Naturally?

Read Naturally®: What Works Clearinghouse
ies.ed.gov/ncee/wWc/interventionreport.aspx?sid=407
Read Naturally was found to have no discernible effects on fluency and reading comprehension.

Will it change to:

"Read Naturally was found to have potentially positive effects on general literacy achievement."

When will the reevaluation of Beginning Reading be completed? Will the Christ study replace Hancock for Beginning Reading? Will the Case 7 Study of First Graders in Forsyth County be evaluated separately? Will Denton, Kemp and Chenault be removed?

Also, I am looking forward to your comments to my recent correspondence (March 5, 2013 – See 3 attachments).

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally® CELEBRATING 20 YEARS

WhatWorks

From:WhatWorks
Sent:21 Mar 2013 14:37:43 +0000
To:'Info@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot (WWCPC 2304)
Dear Mr. Ihnot,

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

What Works Clearinghouse

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

From: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Wednesday, March 20, 2013 4:05 PM
To: What Works
Subject: Read Naturally - Tom Ihnot

Dear Jill Constantine,

Thank you for the courtesy copy of the Adolescent Literacy review of the Heistad study.

Will this result in a change to the following general description that appears when you google Read Naturally?

Read Naturally®: What Works Clearinghouse
ies.ed.gov/ncee/wWc/interventionreport.aspx?sid=407
Read Naturally was found to have no discernible effects on fluency and reading comprehension.

Will it change to:

"Read Naturally was found to have potentially positive effects on general literacy achievement."

When will the reevaluation of Beginning Reading be completed? Will the Christ study replace Hancock for Beginning Reading? Will the Case 7 Study of First Graders in Forsyth County be evaluated separately? Will Denton, Kemp and Chenault be removed?

Also, I am looking forward to your comments to my recent correspondence (March 5, 2013 – See 3 attachments).

Sincerely,

Tom Ihnot

CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: info@readnaturally.com
Sent: Monday, April 08, 2013 4:27 PM
To: What Works
Subject: Tom Ihnot - Read Naturally
Attachments: 2ndConstantine.march.2013.pdf; Addendum 1.pdf; Addendum 2.pdf

Dear Jill Constantine,

I would appreciate your help.

As you know when an educator googles "Read Naturally", the second search result is
What Works Clearinghouse's (WWC) statement: "Read Naturally was found to have no
discernible effects on fluency and reading comphrension".

This is a false statement that continues to cause Read Naturally serious harm. I urge you
to make the change as soon as possible.

The change should be made as a result of WWC's posting of the Heistad study in which
WWC states that "Read Naturally was found to have potentially positive effects". Also,
the WWC improvement index for the Hancock study show an average of +8 percentile
points for fluency.

Please change WWC's statement "no discernible effect" to "Read Naturally showed a
statistically significant positive effect".

In addition, when is the reevaluation of the Beginning Reading section due to be
completed? You should note that Heistad did a separate analysis of the third grade
student for his study. (Case 9 Third Grade Students -
http://www.readnaturally.com/approach/case9.htm)

For more background on these issues, see attachment of my March5, 2013 email to you.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com

www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:10 Apr 2013 13:51:37 +0000
To:'Info@readnaturally.com'
Subject:RE: Tom Ihnot - Read Naturally (WWCPC 2304)
Dear Mr. Ihnot,

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

What Works Clearinghouse

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

From: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Monday, April 08, 2013 4:27 PM
To: What Works
Subject: Tom Ihnot - Read Naturally

Dear Jill Constantine,

I would appreciate your help.

As you know when an educator googles “Read Naturally”, the second search result is What Works Clearinghouse’s (WWC) statement: “Read Naturally was found to have no discernible effects on fluency and reading comphrension”.

This is a false statement that continues to cause Read Naturally serious harm. I urge you to make the change as soon as possible.

The change should be made as a result of WWC’s posting of the Heistad study in which WWC states that “Read Naturally was found to have potentially positive effects”. Also, the WWC improvement index for the Hancock study show an average of +8 percentile points for fluency.

Please change WWC’s statement “no discernible effect” to “Read Naturally showed a statistically significant positive effect”.

In addition, when is the reevaluation of the Beginning Reading section due to be completed? You should note that Heistad did a separate analysis of the third grade student for his study (Case 9 Third Grade Students - http://www.readnaturally.com/approach/case9.htm)

For more background on these issues, see attachment of my March5, 2013 email to you.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally® CELEBRATING 20 YEARS

Mail
www.google.com
Read Naturally Product Catalog 2013
Read Naturally - Better Tools. Better Readers. Brighter Futures.
Read Naturally | National Center on Intensive Intervention
Read Naturally | National Center on Intensive Intervention
Read Naturally® What Works Clearinghouse

info@readnaturally.com

From:info@readnaturally.com
Sent:19 Apr 2013 14:51:49 +0000
To:What Works
Subject:Tom Ihnot
Attachments:NCII.RN.googlesearch.doc
Dear Jill Constantine,

As you know, WWC's statement, "Read Naturally was found to have no discernible effect on fluency and reading comprehension," has greatly damaged Read Naturally, Inc. Not only is the presence of this statement on the WWC website harmful, but the statement is appearing in other places as well.

The National Center on Intensive Intervention (NCII) report about Read Naturally programs contains WWC's statement. WWC's statement also appears as a blurb in a Google search using the search terms NCII and Read Naturally. A basic Google search of Read Naturally also yields WWC's statement as one of the first results. (See attachment)

It is critical that WWC change its statement about Read Naturally, Inc. The statement misinforms educators, damages Read Naturally, Inc., and undermines the NCII. The NCII has posted three reviews of studies of the Read Naturally strategy in which the study authors intended to evaluate the Read Naturally strategy and implemented the strategy correctly. In each study, the NCII reports the "significant positive effects" of the Read Naturally strategy. Click here for more information.

I also urge you to consider that WWC has posted a review of the Heistad study, which shows that the Read Naturally strategy has "statistically significant positive effects."

For six years, Read Naturally, Inc. has suffered from WWC's damaging statement. It is wrong for the company to continue to be harmed by it, and it is a disservice to the educators we both aim to serve.

Please correct this serious problem as soon as possible. I look forward to a quick positive resolution. Please contact me with the timeline to change or remove WWC's conclusion about Read Naturally.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally®
CELEBRATING
20 YEARS

WhatWorks

From:WhatWorks
Sent:19 Apr 2013 20:54:26 +0000
To:'Info@readnaturally.com'
Subject:RE: Tom Ihnot WWCPC 2304
Dear Mr. Ihnot,

Thank you for your email. We are working on a response to your requests.

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: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Friday, April 19, 2013 10:52 AM
To: What Works
Subject: Tom Ihnot

Dear Jill Constantine,

As you know, WWC's statement, "Read Naturally was found to have no discernible effect on fluency and reading comprehension," has greatly damaged Read Naturally, Inc. Not only is the presence of this statement on the WWC website harmful, but the statement is appearing in other places as well.

The National Center on Intensive Intervention (NCII) report about Read Naturally programs contains WWC's statement. WWC's statement also appears as a blurb in a Google search using the search terms NCII and Read Naturally. A basic Google search of Read Naturally also yields WWC's statement as one of the first results. (See attachment)

It is critical that WWC change its statement about Read Naturally, Inc. The statement misinforms educators, damages Read Naturally, Inc., and undermines the NCII. The NCII has posted three reviews of studies of the Read Naturally strategy in which the study authors intended to evaluate the Read Naturally strategy and implemented the strategy correctly. In each study, the NCII reports the "significant positive effects" of the Read Naturally strategy. Click here for more information.

I also urge you to consider that WWC has posted a review of the Heistad study, which shows that the Read Naturally strategy has "statistically significant positive effects."

For six years, Read Naturally, Inc. has suffered from WWC's damaging statement. It is wrong for the company to continue to be harmed by it, and it is a disservice to the educators we both aim to serve.

Please correct this serious problem as soon as possible. I look forward to a quick positive resolution. Please contact me with the timeline to change or remove WWC's conclusion about Read Naturally.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: info@readnaturally.com
Sent: Thursday, May 02, 2013 3:11 PM
To: What Works
Subject: Read Naturally

Dear Jill Constantine,

Thank you for your continued consideration of WWC's reviews of the Read Naturally strategy. The new reviews of the Read Naturally strategy by WWC and the NCII over the past six months have provided educators with important information about our solutions for struggling readers.

Unfortunately, a google search of "NCII – Read Naturally" still produces WWC's statement, "Read Naturally was found to have no discernible effects on reading fluency and comprehension" instead of the positive effects NCII noted in its review of the Heistad, Christ and Tucker studies. Furthermore, if an educator simply googles "Read Naturally," WWC's negative statement continues to be the second search result in spite of WWC's recent positive review based on the Heistad study.

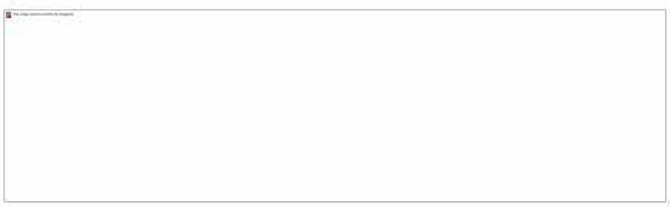
I've explained in great detail in prior emails why WWC's negative statement about Read Naturally, Inc. is misleading. In addition, please consider that a significant percentage of the participants receiving the Read Naturally strategy in the Heistad study were third graders and would thus fit in the Beginning Reading section of WWC reviews. The positive effects noted for this population provide another reason why WWC's negative statement about Read Naturally, Inc. in its Beginning Reading section is misleading.

In light of the positive NCII review and WWC's positive review of the Heistad study, it is time to fix WWC's negative statement about Read Naturally, Inc. The high visibility of this statement in web searches misinforms educators and wrongly damages Read Naturally, Inc. I would be happy to consider your point of view and together create a statement that upholds our mutual commitment to providing educators with the accurate information they need to do their jobs effectively. Please respond and let me know how we can proceed.

Sincerely,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax



Julia Lyskawa

From:Julia Lyskawa
Sent:13 May 2013 15:37:01 -0400
To:Loyco, Kate;Allison Young
Subject:FW: What Works Clearinghouse (QR2010013)
Attachments:QRT2010013_Response 5.13.2013.pdf
FYI – The response to Tom Ihnot went out today.

How did it go last week while I was out? Any major issues or new QRT-related emails?

From: What Works
Sent: Monday, May 13, 2013 3:32 PM
To: 'kramy@readnaturally.com'; 'info@readnaturally.com'
Subject: What Works Clearinghouse (QR2010013)

Dear Mr. Ihnot:

Attached is a response to the questions you raised in your inquiries concerning the What Works Clearinghouse reviews of studies of Read Naturally®.

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/.

May 12, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013, WWC2304, and WWCPC3661

Mr. Ihnot:

Thank you for your response to Dr. Constantine's letter of February 26, 2013. The What Works Clearinghouse (WWC) Quality Review Team has reviewed your correspondence dated March 5, March 6, and March 20, 2013.

Your correspondence raises five questions with respect to reviews of Read Naturally®: (1) results when the term "Read Naturally" is searched using Google; (2) the inclusion of five studies (Hancock 2002; Kemp 2006; Denton et al. 2006; Chenault 2004; and Chenault et al. 2006); (3) the update to the Beginning Reading intervention report on Read Naturally®; (4) reviews by the National Center on Intensive Intervention (NCII) of studies of Read Naturally®; and (5) the difficulties an individual might have in distinguishing between Read Naturally, Inc. and Read Naturally®. Below we provide responses to these questions, including a brief summary of our reasoning in each area. The documents explaining all of our procedures and review criteria continue to be publically available on the WWC website. As suggested in our previous letters, please refer to the *WWC Procedures and Standards Handbook* and the topic area review protocol for a particular intervention report.

Google Results for "Read Naturally"

WWC does not have any control over or input as to what results are presented by Google or any other search engine when an individual searches "Read Naturally." It is our understanding that the various search engines use algorithms that generally are proprietary.

Inclusion of Hancock (2002), Kemp (2006), Denton et al. (2006), Chenault (2004), and Chenault et al. (2006)

We appreciate the information you provided to us in Addendums 1 and 2 of your March correspondence. Our response is the same as in our February 26 letter: WWC followed protocol in determining whether studies supported an examination of the effectiveness of Read Naturally®. To assist the reader in understanding what was being examined, WWC includes in the intervention report a description of both the intervention and comparison conditions to the extent they are described in each individual study. This includes describing whether elements of an intervention were not implemented. WWC reviews publicly available research, and these reports may differ in the amount of detail provided on the condition of interest to WWC. To the extent possible, the reports summarize what is known about how a particular intervention was implemented.

Update to the Beginning Reading Intervention Report on Read Naturally®

WWC is currently updating the Beginning Reading intervention report for Read Naturally®. As the developer and distributor of Read Naturally®, a courtesy copy of the embargoed intervention report will be shared with you 24 hours before release. We cannot provide details on the content of the update before you receive the copy, which is embargoed until it is released. All of the studies you listed in your March 20 email were reviewed as part of the updated report.

NCII Reviews of Read Naturally® Studies

Eight of the nine studies that you identified in Addendums 1 and 2 were reviewed for the update to the Beginning Reading intervention report on Read Naturally® (Hesitad, Christ, Tucker, and five case studies). As the report has not yet been released, we cannot discuss individual study ratings or the overall summary of effectiveness. The remaining case study (Case 5, "Throwing Sixth Graders a Lifeline") is not within the age range of the Beginning Reading review and thus is not eligible for review in a Beginning Reading report, although it could be eligible for review in the Adolescent Literacy report on Read Naturally®.

Distinguishing Between Read Naturally, Inc. and Read Naturally®

The program details, which are included in the main part of the intervention report, are shared with developers or distributors for review. We provided you with that description, and you provided input in February 2012. However, we agree that it is important and useful to readers to distinguish among the different products that support Read Naturally® strategies. Thus, we will revise the product description further to discuss Read Naturally® as a strategy supported by the four products (Read Naturally® Live, Read Naturally® Encore, Read Naturally® ME, and Read Naturally® SE) that can be purchased from Read Naturally®, Inc. To the extent that the specific products are described in each study, we will also clearly note which products were used. In addition, we will use the copyright symbol (®) with the phrase Read Naturally® throughout the study.

WWC appreciates your careful review of our products and processes. If you or another member of Read Naturally, Inc. are interested in learning more about WWC's review standards, please consider applying for one of the training sessions offered by the Institute of Education Sciences. You can send an email to the WWC help desk asking to be added to the email list for training announcements.

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

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

From: info@readnaturally.com
Sent: Wednesday, May 15, 2013 2:07 PM
To: What Works
Subject: Jill Constantine

Dear Jill Constantine,

Thank you for your response. I look forward to more accurate reports on Read Naturally, Inc.
products when you update your Beginning Reading and Adolescent Literature sections.

In the meantime, it is very important to correct the statement, "Read Naturally was found to have
no discernible effects on fluency and reading comprehension." I understand that WWC has no
influence over Google search formulas. However, WWC does control the wording of this
statement.

This statement is obsolete for the following reasons:

1. WWC recently posted a review of the Heistad Four School Study where WWC states that
a Read Naturally, Inc. product showed a "statistically significant positive effect." This
review is overridden by the "no discernible effect" statement.
2. NCII is also under the United States Department of Education, and its reviews clearly
state that the Read Naturally © strategy has positive effects. The WWC statement needs
to be updated to reflect the evolving process of reviews.
3. In the original WWC Beginning Reading review, the Mesa study was inexplicably grouped
with the Hancock study. The Mesa study should stand on its own. It clearly demonstrates
the Read Naturally © strategy's significant positive effects in fluency and comprehension.
4. As you stated in your email, it is important to distinguish between Read Naturally, Inc.
and the products used in the studies. Simply saying "Read Naturally was found…" has
no benefit to educators as they are looking for specific products, and our company
offers dozens of solutions targeting different populations.
5. In both the Hancock and Kemp studies, the control and treatment groups had similar
fluency gains and in both studies those fluency gains were greater than average growth
based on oral reading fluency norms. It is therefore inaccurate to state that the students
who used Read Naturally passages had no discernible effects on fluency.
6. When the updated Beginning Reading intervention report is updated, the Christ Study,

Heistad Study (3rd grade), Mesa Study and any other control group case studies you choose to use will clearly demonstrate that the Read Naturally © strategy has discernible effects on fluency, rate, accuracy and comprehension. The current statement provides educators with an inconsistent message from WWC.

7. Educators typically do not have the time to analyze all the details in the intervention report on implementation. They see the headline. Now that WWC and NCII have posted new reviews on Read Naturally, Inc. products, it is important for educators to have the most up-to-date headline information.

Please change WWC's statement about Read Naturally, Inc. and our products as soon as possible. For the reasons stated above, not changing the statement in favor of waiting for the updated reviews will continue to propagate misleading information and damage Read Naturally, Inc. for an unnecessary period of time.

Concerning the Hancock, Denton, Kemp and Chenault studies, I must reiterate that there must have been something wrong with the protocol. In each of the studies, the authors have been very clear:

Hancock:
"While I used Read Naturally materials, I did NOT fully implement the Read Naturally strategy and my study was NOT intended to evaluate the Read Naturally strategy." -Carrie Hancock, Study Author

Denton:
"The measures used to evaluate the effectiveness of the instruction did not include measures of fluency. Since Read Naturally is an intervention targeted at fluency, the measures used did not match the specific purpose and intent of the Read Naturally instruction given to the students in that group. The Denton study is misrepresented and erroneously mislabeled as an evaluation study." - Jan Hasbrouck, Study Co-Author

Kemp:
"My study was a study of sustained silent reading and the purpose was not to evaluate Read Naturally." - Susan Kemp, Study Author

"Another limitation [of the study] is that Read Naturally was not just used with at-risk readers. The Read Naturally program was developed to help students who are at-risk. Typically this would represent students who are scoring well below the 25th percentile. Students participating in this study did not fit the description of at-risk. Many scored well above the 25th percentile." – Susan Kemp, page 74 of dissertation

Chenault:

"This study was not intended as an examination of the Read Naturally program. The lack of demonstrable standard score gains on the GORT Rate subtest for the Read Naturally group following this short term of training should not be construed as indicating that this instructional program is not effective in improving reading fluency. These children were not selected for their reading fluency problems – they were selected for their persisting writing problems." – Belle Chenault, Study Author, dissertation

"The Read Naturally program is a widely recognized and widely used curriculum that has been appropriately validated in other studies. This study was not intended as an examination of the Read Naturally program. Using our study to evaluate Read Naturally or its effects is a misapplication of our data." – Belle Chenault

"You cannot use a contact control (which we knew would be good for the children and not a waste of their time but was not predicted to change the aspect of writing we were training) to draw conclusions about the effectiveness of the contact control. In fact there is an emerging literature on the fact that, while reading and writing have some relationships, there are aspects of writing not related to reading at all, which is why I think Joe Jenkins proposed using reading fluency training [Read Naturally, Inc. product] as a contact control."
– Dr. Virginia Berninger, Belle Chenault's advisor

"In no way can this kind of experimental design, which is well established in educational science, be considered a randomized controlled evaluation of the contact control for its treatment effectiveness." – Dr. Virginia Berninger

Thank you for inviting me to learn more about WWC's review processes. Please add info@readnaturally.com to your email list for training announcements.

I am pleased with the opportunity to communicate and clarify the issues concerning Read Naturally, Inc. and WWC's statement about our company and products. Changing the current WWC statement immediately is critically important for educators and Read Naturally, Inc. I appreciate your willingness to work with me on this issue and look forward to a resolution that upholds the mission of WWC, fairly represents Read Naturally, Inc., and provides educators with the accurate information they deserve.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:30 May 2013 19:42:02 +0000
To:'Info@readnaturally.com'
Cc:'NSeftor@mathematica-mpr.com'
Subject:What Works Clearinghouse (QR2010013)
Attachments:QRT2010013 February 26 2013.pdf, QRT2010013 May 12 2013.pdf

Dear Mr. Ihnot,

The WWC is working on the update to the Beginning Reading Intervention Report on Read Naturally®, including revisiting the program description and program details to clarify the differences between Read Naturally, Inc., the Read Naturally strategy, and the four Read Naturally® products (ME, SE, Encore and Live). You will receive an embargoed copy of the report 24 hours before publication. Once released, the old report will be replaced on the web and the searches will report the findings in the updated report.

Thank you for continuing to bring to our attention your concerns about the studies reviewed by the WWC for the Read Naturally® intervention reports. We have previously responded to your concerns about the NCII reviews and the various studies (Mesa, Hancock, Kemp, Christ, Heistad, Denton, Kemp and Chenault). Please see our correspondence to you in February 2013 and May 2013, attached. If you have any different concerns, please let us know.

What Works Clearinghouse

February 26, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013

Dear Mr. Ihnot:

Thank you for your inquiries concerning the What Works Clearinghouse (WWC) reviews of four studies of Read Naturally®—Kemp (2006), Denton et al. (2004), Hancock (2002), and Chenault et al. (2006)—and for your email regarding my response to Dr. Kevin Feldman's inquiry about the WWC reviews of studies of Read Naturally®. In response to your concerns, we conducted an independent quality review. The WWC quality review team responds to concerns raised by study authors, curriculum developers, or other relevant parties about WWC reviews published on our website. When a quality review is conducted, a researcher who was not involved in the initial reviews undertakes an independent assessment of the studies in question. The researcher also investigates the procedures used and decisions made during the original reviews of the studies. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the WWC.

The quality review considered the concerns from your emails regarding study inclusion, implementation, student sample, and outcomes.

Study Inclusion

The WWC followed protocol in choosing to review these four studies. The WWC screens studies based on a number of factors, including relevancy and methodology criteria (as described in the WWC Procedures and Standards Handbook, pages 8–10, available at http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19). The WWC does not screen based on whether the author(s) explicitly intended the study as an evaluation of an intervention, but rather whether the study presents a primary analysis of the effect of an intervention. This screening process allows for the review of a wide range of relevant and methodologically sound studies.

Thank you for bringing to my attention the studies by Mesa (2004) and Tucker (2010). The Tucker (2010) study will be included in our literature search for updating the Intervention Report

for Read Naturally®, reviewed under the Beginning Reading Evidence Review Protocol. In addition, as you note, the Mesa (2004) study was included in the original review of Read Naturally®, but will be reviewed again in an updated review.

Implementation

The WWC followed protocol in the manner in which the four studies are described in WWC publications. The WWC Handbook says, "The WWC makes no adjustments or corrections for variations in implementation of the intervention; however, if a study meets standards and is included in an intervention report, descriptions of implementation are provided in the report appendices to provide context for the findings" (page 16). This approach is appropriate because there is no standard metric for fidelity to intervention design. Thus, the WWC includes studies with variation in fidelity and does not evaluate implementation fidelity. The quality review team verified that variations in implementation that are noted in the four studies and that may affect the interpretation of findings were properly included in the WWC publications.

Regarding the Kemp (2006) study, your emails raised a concern about implementation of the Read Naturally® program. The quality review noted that the study does not provide any indication of deviation from program design. Specifically, the study states, "... it could be concluded that all teachers implemented the Read Naturally® program as prescribed" (page 40). Based on this information, the quality review concluded that there was no evidence that deviations in implementation should have been noted in the WWC Intervention Report. Regarding your specific concern about the teacher-to-student ratio, the study does not note any deviation from the Read Naturally® program in this area and does not provide any information on the ratio.

You also expressed concern over deviations in the implementation of Read Naturally® in the Denton et al. (2004) study. The quality review found that the inclusion of additional activities was noted in the study and reported in the WWC Intervention Report (see footnote 7 on page 3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). The other potential deviations that you noted (accountability to pass criteria, use of audio tapes, and the number of sessions per week) were not identified in the study and are therefore not noted in the WWC publication.

Regarding your concerns that the Read Naturally® strategy was not implemented fully for the Hancock (2002) study, the quality review found that WWC procedures were followed in the WWC Intervention Report. The WWC contacted Hancock regarding the implementation of Read Naturally® in the 2002 study. Her response indicated that the study excluded Read Naturally®'s pre-reading vocabulary instruction component and the placement system to individualize instruction. The WWC Intervention Report was revised to note these deviations in implementation (see footnote 4 on page 2, available at http://ies.ed.gov/ncee/wwc/InterventionReport.aspx?sid=407). Other potential concerns with implementation described in your email were not reported by the author and were not included in the revised WWC Intervention Report.

The quality review also investigated concerns that you raised in prior emails. The email from June 2012 regarding Dr. Feldman's questions was addressed in my letter to Dr. Feldman; I will not

repeat the quality review findings here. Your email from May 2012 included some text that was reported to be a quote from Hancock's email to the WWC, but the quoted statements were not part of Hancock's email text. The misreported quote is also found on the Read Naturally® website at http://www.readnaturally.com/approach/hancock.htm.

Regarding the Chenault et al. (2006) study, you expressed concern with the implementation of Read Naturally® related to the selection of students and the limited number of sessions. The quality review found that the WWC Intervention Report appropriately notes both of these factors (see pages 2–3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=409).

Student Sample

You also raised a concern that the Kemp (2006) study included a large share of students who were above the 50th percentile in reading fluency who were not "at-risk" readers. The quality review verified that the WWC Intervention Report accurately states that the study population for the results reported by the WWC was third-grade English language learner (ELL) students (page 2, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). Furthermore, the quality review noted that Read Naturally® program materials state that the program "… products help students of all ages and abilities—whether special education, ELL, Title I, or mainstream—to improve their reading skills, gain confidence, and transform their attitudes toward school and learning" (available at http://www.readnaturally.com/products/default.htm).

Finally, regarding the differences in pre-test performance among native-speaking students in the treatment and control groups, the quality review found that these concerns are not relevant to this WWC publication, which reports on findings only for English language learner students. Furthermore, the quality review concluded that the WWC appropriately restricted the reported findings to impacts assessed relative to the comparison group, and thus did not include achievement gains in oral reading fluency.

Outcomes

Regarding the Denton et al. (2004) study, you expressed concern that the study did not assess fluency. The quality review noted that the WWC Intervention Report clearly states that the study assessed reading achievement (page 4, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408), an outcome of relevance to the WWC ELL topic area. The quality review noted that the Read Naturally® program materials state there are several components of the program beyond fluency that are related to reading achievement: "Our industry-leading programs develop and support the five essential components of reading, identified by the National Reading Panel: phonemic awareness, phonics, fluency, vocabulary, and comprehension" (see "The Read Naturally Strategy," available at http://www.readnaturally.com/approach/default.htm).

I hope that this letter has addressed your concerns. You requested that we meet to discuss the WWC reviews of Read Naturally®. However, WWC policy requires all correspondence in writing. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov.

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

cc: Joy Lesnick, IES

May 12, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013, WWC2304, and WWCPC3661

Mr. Ihnot:

Thank you for your response to Dr. Constantine's letter of February 26, 2013. The What Works Clearinghouse (WWC) Quality Review Team has reviewed your correspondence dated March 5, March 6, and March 20, 2013.

Your correspondence raises five questions with respect to reviews of Read Naturally®: (1) results when the term "Read Naturally" is searched using Google; (2) the inclusion of five studies (Hancock 2002; Kemp 2006; Denton et al. 2006; Chenault 2004; and Chenault et al. 2006); (3) the update to the Beginning Reading intervention report on Read Naturally®; (4) reviews by the National Center on Intensive Intervention (NCII) of studies of Read Naturally®; and (5) the difficulties an individual might have in distinguishing between Read Naturally, Inc. and Read Naturally®. Below we provide responses to these questions, including a brief summary of our reasoning in each area. The documents explaining all of our procedures and review criteria continue to be publically available on the WWC website. As suggested in our previous letters, please refer to the *WWC Procedures and Standards Handbook* and the topic area review protocol for a particular intervention report.

Google Results for "Read Naturally"

WWC does not have any control over or input as to what results are presented by Google or any other search engine when an individual searches "Read Naturally." It is our understanding that the various search engines use algorithms that generally are proprietary.

Inclusion of Hancock (2002), Kemp (2006), Denton et al. (2006), Chenault (2004), and Chenault et al. (2006)

We appreciate the information you provided to us in Addendums 1 and 2 of your March correspondence. Our response is the same as in our February 26 letter: WWC followed protocol in determining whether studies supported an examination of the effectiveness of Read Naturally®. To assist the reader in understanding what was being examined, WWC includes in the intervention report a description of both the intervention and comparison conditions to the extent they are described in each individual study. This includes describing whether elements of an intervention were not implemented. WWC reviews publicly available research, and these reports may differ in the amount of detail provided on the condition of interest to WWC. To the extent possible, the reports summarize what is known about how a particular intervention was implemented.

Update to the Beginning Reading Intervention Report on Read Naturally®

WWC is currently updating the Beginning Reading intervention report for Read Naturally®. As the developer and distributor of Read Naturally®, a courtesy copy of the embargoed intervention report will be shared with you 24 hours before release. We cannot provide details on the content of the update before you receive the copy, which is embargoed until it is released. All of the studies you listed in your March 20 email were reviewed as part of the updated report.

NCII Reviews of Read Naturally® Studies

Eight of the nine studies that you identified in Addendums 1 and 2 were reviewed for the update to the Beginning Reading intervention report on Read Naturally® (Hesitad, Christ, Tucker, and five case studies). As the report has not yet been released, we cannot discuss individual study ratings or the overall summary of effectiveness. The remaining case study (Case 5, "Throwing Sixth Graders a Lifeline") is not within the age range of the Beginning Reading review and thus is not eligible for review in a Beginning Reading report, although it could be eligible for review in the Adolescent Literacy report on Read Naturally®.

Distinguishing Between Read Naturally, Inc. and Read Naturally®

The program details, which are included in the main part of the intervention report, are shared with developers or distributors for review. We provided you with that description, and you provided input in February 2012. However, we agree that it is important and useful to readers to distinguish among the different products that support Read Naturally® strategies. Thus, we will revise the product description further to discuss Read Naturally® as a strategy supported by the four products (Read Naturally® Live, Read Naturally® Encore, Read Naturally® ME, and Read Naturally® SE) that can be purchased from Read Naturally®, Inc. To the extent that the specific products are described in each study, we will also clearly note which products were used. In addition, we will use the copyright symbol (®) with the phrase Read Naturally® throughout the study.

WWC appreciates your careful review of our products and processes. If you or another member of Read Naturally, Inc. are interested in learning more about WWC's review standards, please consider applying for one of the training sessions offered by the Institute of Education Sciences. You can send an email to the WWC help desk asking to be added to the email list for training announcements.

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

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

February 26, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013

Dear Mr. Ihnot:

Thank you for your inquiries concerning the What Works Clearinghouse (WWC) reviews of four studies of Read Naturally®—Kemp (2006), Denton et al. (2004), Hancock (2002), and Chenault et al. (2006)—and for your email regarding my response to Dr. Kevin Feldman's inquiry about the WWC reviews of studies of Read Naturally®. In response to your concerns, we conducted an independent quality review. The WWC quality review team responds to concerns raised by study authors, curriculum developers, or other relevant parties about WWC reviews published on our website. When a quality review is conducted, a researcher who was not involved in the initial reviews undertakes an independent assessment of the studies in question. The researcher also investigates the procedures used and decisions made during the original reviews of the studies. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the WWC.

The quality review considered the concerns from your emails regarding study inclusion, implementation, student sample, and outcomes.

Study Inclusion

The WWC followed protocol in choosing to review these four studies. The WWC screens studies based on a number of factors, including relevancy and methodology criteria (as described in the WWC Procedures and Standards Handbook, pages 8–10, available at http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19). The WWC does not screen based on whether the author(s) explicitly intended the study as an evaluation of an intervention, but rather whether the study presents a primary analysis of the effect of an intervention. This screening process allows for the review of a wide range of relevant and methodologically sound studies.

Thank you for bringing to my attention the studies by Mesa (2004) and Tucker (2010). The Tucker (2010) study will be included in our literature search for updating the Intervention Report

for Read Naturally®, reviewed under the Beginning Reading Evidence Review Protocol. In addition, as you note, the Mesa (2004) study was included in the original review of Read Naturally®, but will be reviewed again in an updated review.

Implementation

The WWC followed protocol in the manner in which the four studies are described in WWC publications. The WWC Handbook says, "The WWC makes no adjustments or corrections for variations in implementation of the intervention; however, if a study meets standards and is included in an intervention report, descriptions of implementation are provided in the report appendices to provide context for the findings" (page 16). This approach is appropriate because there is no standard metric for fidelity to intervention design. Thus, the WWC includes studies with variation in fidelity and does not evaluate implementation fidelity. The quality review team verified that variations in implementation that are noted in the four studies and that may affect the interpretation of findings were properly included in the WWC publications.

Regarding the Kemp (2006) study, your emails raised a concern about implementation of the Read Naturally® program. The quality review noted that the study does not provide any indication of deviation from program design. Specifically, the study states, "… it could be concluded that all teachers implemented the Read Naturally® program as prescribed" (page 40). Based on this information, the quality review concluded that there was no evidence that deviations in implementation should have been noted in the WWC Intervention Report. Regarding your specific concern about the teacher-to-student ratio, the study does not note any deviation from the Read Naturally® program in this area and does not provide any information on the ratio.

You also expressed concern over deviations in the implementation of Read Naturally® in the Denton et al. (2004) study. The quality review found that the inclusion of additional activities was noted in the study and reported in the WWC Intervention Report (see footnote 7 on page 3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). The other potential deviations that you noted (accountability to pass criteria, use of audio tapes, and the number of sessions per week) were not identified in the study and are therefore not noted in the WWC publication.

Regarding your concerns that the Read Naturally® strategy was not implemented fully for the Hancock (2002) study, the quality review found that WWC procedures were followed in the WWC Intervention Report. The WWC contacted Hancock regarding the implementation of Read Naturally® in the 2002 study. Her response indicated that the study excluded Read Naturally®'s pre-reading vocabulary instruction component and the placement system to individualize instruction. The WWC Intervention Report was revised to note these deviations in implementation (see footnote 4 on page 2, available at http://ies.ed.gov/ncee/wwc/InterventionReport.aspx?sid=407). Other potential concerns with implementation described in your email were not reported by the author and were not included in the revised WWC Intervention Report.

The quality review also investigated concerns that you raised in prior emails. The email from June 2012 regarding Dr. Feldman's questions was addressed in my letter to Dr. Feldman; I will not

repeat the quality review findings here. Your email from May 2012 included some text that was reported to be a quote from Hancock's email to the WWC, but the quoted statements were not part of Hancock's email text. The misreported quote is also found on the Read Naturally® website at http://www.readnaturally.com/approach/hancock.htm.

Regarding the Chenault et al. (2006) study, you expressed concern with the implementation of Read Naturally® related to the selection of students and the limited number of sessions. The quality review found that the WWC Intervention Report appropriately notes both of these factors (see pages 2–3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=409).

Student Sample

You also raised a concern that the Kemp (2006) study included a large share of students who were above the 50th percentile in reading fluency who were not "at-risk" readers. The quality review verified that the WWC Intervention Report accurately states that the study population for the results reported by the WWC was third-grade English language learner (ELL) students (page 2, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). Furthermore, the quality review noted that Read Naturally® program materials state that the program "… products help students of all ages and abilities—whether special education, ELL, Title I, or mainstream—to improve their reading skills, gain confidence, and transform their attitudes toward school and learning" (available at http://www.readnaturally.com/products/default.htm).

Finally, regarding the differences in pre-test performance among native-speaking students in the treatment and control groups, the quality review found that these concerns are not relevant to this WWC publication, which reports on findings only for English language learner students. Furthermore, the quality review concluded that the WWC appropriately restricted the reported findings to impacts assessed relative to the comparison group, and thus did not include achievement gains in oral reading fluency.

Outcomes

Regarding the Denton et al. (2004) study, you expressed concern that the study did not assess fluency. The quality review noted that the WWC Intervention Report clearly states that the study assessed reading achievement (page 4, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408), an outcome of relevance to the WWC ELL topic area. The quality review noted that the Read Naturally® program materials state there are several components of the program beyond fluency that are related to reading achievement: "Our industry-leading programs develop and support the five essential components of reading, identified by the National Reading Panel: phonemic awareness, phonics, fluency, vocabulary, and comprehension" (see "The Read Naturally Strategy," available at http://www.readnaturally.com/approach/default.htm).

I hope that this letter has addressed your concerns. You requested that we meet to discuss the WWC reviews of Read Naturally®. However, WWC policy requires all correspondence in writing. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov.

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

cc: Joy Lesnick, IES

info@readnaturally.com

From:info@readnaturally.com
Sent:3 Jun 2013 13:41:07 +0000
To:What Works
Subject:Read Naturally - Tom Ihnot
Attachments:QRT2010013 February 26 2013.pdf, QRT2010013 May 12 2013.pdf, Addendum 1.docx

Dear Jill Constantine,

Thank you for the update. I appreciate that the WWC Quality Review Team has researched the issues related to WWC's reviews of Read Naturally, Inc. and Read Naturally® programs. I am grateful that you will be replacing the statement, "Read Naturally was found to have no discernible effects on reading fluency and comprehension." Read Naturally, Inc. has endured this inaccurate and damaging statement since 2006. I am excited that WWC will soon provide educators with an accurate description of Read Naturally® programs based on the Mesa, Christ, Heistad, Tucker and other studies.

Thank you for addressing many of my concerns in your emails. Your email from May 12, 2013 email does not address many of the issues I raised in Addendum 1 (attached) of my email to you on March 5. I still agree with Hancock, Hasbrouck (Denton Study), Kemp and Chenault that their studies should not be used as reviews of Read Naturally® programs as the studies were not primary analyses of the effect of an intervention. Addendum 1 spells out additional implementation variations. I have attached Addendum 1 to this email and hope you will review it.

However, the vast majority of educators will make judgments based on WWC's headline description. Thus, I am extremely happy that WWC will soon replace its inaccurate statement about Read Naturally, Inc. and Read Naturally® programs. I look forward to the revised report and updated statement.

Sincerely,

Tom Ihnot
CEO
Read Naturally, Inc.

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: WhatWorks [mailto:What.Works@icfi.com]
Sent: Thursday, May 30, 2013 2:42 PM
To: info@readnaturally.com
Cc: NSeftor@mathematica-mpr.com
Subject: What Works Clearinghouse (QR2010013)

Dear Mr. Ihnot,

The WWC is working on the update to the Beginning Reading Intervention Report on Read Naturally®, including revisiting the program description and program details to clarify the differences between Read Naturally, Inc., the Read Naturally strategy, and the four Read Naturally® products (ME, SE, Encore and Live). You will receive an embargoed copy of the report 24 hours before publication. Once released, the old report will be replaced on the web and the searches will report the findings in the updated report.

Thank you for continuing to bring to our attention your concerns about the studies reviewed by the WWC for the Read Naturally® intervention reports. We have previously responded to your concerns about the NCII reviews and the various studies (Mesa, Hancock, Kemp, Christ, Heistad, Denton, Kemp and Chenault). Please see our correspondence to you in February 2013 and May 2013, attached. If you have any different concerns, please let us know.

What Works Clearinghouse

Addendum 1

The passages published by Read Naturally, Inc. are an attractive tool for researchers because they are carefully leveled short passages at a wide range of reading levels. The Hancock, Kemp, Denton, and Chenault studies all used passages published by Read Naturally, Inc., but none used the trademarked Read Naturally®, strategy. Therefore, they cannot be used to evaluate the Read Naturally strategy or a specific Read Naturally product.

Hancock Study

In addressing the Hancock study, WWC states in the footnotes that Hancock indicated her study excluded the Read Naturally strategy's pre-reading vocabulary instruction and the placement system to individualize instruction. These categories represent a wide range of major deviations, and there are more that should be included:

- Prediction step
- Adjusting goals and levels
- Accountability for passing

Carrie Hancock states in her dissertation that critical aspects of the Read Naturally strategy were not part of her study and that all aspects of the Read Naturally strategy should be included in fluency interventions.

Consequently, since Hancock did not implement the Read Naturally strategy, her study cannot be used to evaluate the Read Naturally strategy as a primary analysis of the effect of an intervention. The details of the Hancock Study Deviations are described in the attached copy of a previous email. Another attachment verifies the quote from Carrie Hancock that is on the Read Naturally website.

Carrie Hancock can be contacted at (b)(6) to verify these deviations and that she only used passages published by Read Naturally, Inc. and not the Read Naturally strategy.

Kemp Study

Concerning the Kemp study, WWC quotes the study's statement: "It could be concluded that all teachers implemented the Read Naturally program as prescribed." This statement means that Kemp used passages published by Read Naturally, Inc. in her model. However, her model was not the Read Naturally strategy.

On page 74 of her dissertation, Kemp clearly states that her study was not appropriate for implementing the Read Naturally strategy since the vast majority of the students were not at risk. Of course, she did not implement the Read Naturally strategy and consequently her study should not be used as a primary analysis of an intervention.

In spite of the Kemp study's design flaws, the group using passages from Read Naturally, Inc. did make significant gains in fluency. These students averaged a gain of 1.5 WCPM (words correct per minute) each week. Oral reading fluency norms show that the average fluency gain per week for third grade students is 1.1 WCPM. This data indicates that the students using the passages from Read Naturally, Inc. were accelerating their progress.

Susan Kemp can be contacted at (b)(6) or (b)(6), and more information can be found here: http://www.readnaturally.com/approach/kemp.htm

Denton Study

Concerning your comments on the Denton study, you cite the statement on our website: "Our industry-leading programs develop and support the five essential components …" The Read Naturally strategy and its flagship programs: Masters Edition, Encore, Software Edition, and Read Live develop fluency, support vocabulary, and promote comprehension. We offer other programs that develop phonics and support phonemic awareness – Group and Tutoring Edition (GATE), Word Warm-ups, and Signs for Sounds. The Denton study did not come close to using the Read Naturally strategy and did not use GATE, Word Warm-ups or Signs for Sounds for phonics or phonemic awareness.

You can contact the Denton study advisor, Dr. Jan Hasbrouck at (b)(6) or 206-498-0123. More information can be found here: http://www.readnaturally.com/approach/denton.htm

Chenault Study

The Chenault study authors have strongly stated their objections for using their study as an evaluation of the Read Naturally strategy. You can contact Belle Chenault at (b)(6) or (b)(6) and study advisor Dr. Virginia Berninger at vwb@u.washington.edu or 206-616-6372. More information can be found here: http://www.readnaturally.com/approach/chenault.htm

May 12, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013, WWC2304, and WWCPC3661

Mr. Ihnot:

Thank you for your response to Dr. Constantine's letter of February 26, 2013. The What Works Clearinghouse (WWC) Quality Review Team has reviewed your correspondence dated March 5, March 6, and March 20, 2013.

Your correspondence raises five questions with respect to reviews of Read Naturally®: (1) results when the term "Read Naturally" is searched using Google; (2) the inclusion of five studies (Hancock 2002; Kemp 2006; Denton et al. 2006; Chenault 2004; and Chenault et al. 2006); (3) the update to the Beginning Reading intervention report on Read Naturally®; (4) reviews by the National Center on Intensive Intervention (NCII) of studies of Read Naturally®; and (5) the difficulties an individual might have in distinguishing between Read Naturally, Inc. and Read Naturally®. Below we provide responses to these questions, including a brief summary of our reasoning in each area. The documents explaining all of our procedures and review criteria continue to be publically available on the WWC website. As suggested in our previous letters, please refer to the *WWC Procedures and Standards Handbook* and the topic area review protocol for a particular intervention report.

Google Results for "Read Naturally"

WWC does not have any control over or input as to what results are presented by Google or any other search engine when an individual searches "Read Naturally." It is our understanding that the various search engines use algorithms that generally are proprietary.

Inclusion of Hancock (2002), Kemp (2006), Denton et al. (2006), Chenault (2004), and Chenault et al. (2006)

We appreciate the information you provided to us in Addendums 1 and 2 of your March correspondence. Our response is the same as in our February 26 letter: WWC followed protocol in determining whether studies supported an examination of the effectiveness of Read Naturally®. To assist the reader in understanding what was being examined, WWC includes in the intervention report a description of both the intervention and comparison conditions to the extent they are described in each individual study. This includes describing whether elements of an intervention were not implemented. WWC reviews publicly available research, and these reports may differ in the amount of detail provided on the condition of interest to WWC. To the extent possible, the reports summarize what is known about how a particular intervention was implemented.

Update to the Beginning Reading Intervention Report on Read Naturally®

WWC is currently updating the Beginning Reading intervention report for Read Naturally®. As the developer and distributor of Read Naturally®, a courtesy copy of the embargoed intervention report will be shared with you 24 hours before release. We cannot provide details on the content of the update before you receive the copy, which is embargoed until it is released. All of the studies you listed in your March 20 email were reviewed as part of the updated report.

NCII Reviews of Read Naturally® Studies

Eight of the nine studies that you identified in Addendums 1 and 2 were reviewed for the update to the Beginning Reading intervention report on Read Naturally® (Hesitad, Christ, Tucker, and five case studies). As the report has not yet been released, we cannot discuss individual study ratings or the overall summary of effectiveness. The remaining case study (Case 5, "Throwing Sixth Graders a Lifeline") is not within the age range of the Beginning Reading review and thus is not eligible for review in a Beginning Reading report, although it could be eligible for review in the Adolescent Literacy report on Read Naturally®.

Distinguishing Between Read Naturally, Inc. and Read Naturally®

The program details, which are included in the main part of the intervention report, are shared with developers or distributors for review. We provided you with that description, and you provided input in February 2012. However, we agree that it is important and useful to readers to distinguish among the different products that support Read Naturally® strategies. Thus, we will revise the product description further to discuss Read Naturally® as a strategy supported by the four products (Read Naturally® Live, Read Naturally® Encore, Read Naturally® ME, and Read Naturally® SE) that can be purchased from Read Naturally®, Inc. To the extent that the specific products are described in each study, we will also clearly note which products were used. In addition, we will use the copyright symbol (®) with the phrase Read Naturally® throughout the study.

WWC appreciates your careful review of our products and processes. If you or another member of Read Naturally, Inc. are interested in learning more about WWC's review standards, please consider applying for one of the training sessions offered by the Institute of Education Sciences. You can send an email to the WWC help desk asking to be added to the email list for training announcements.

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

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

WhatWorks

From:WhatWorks
Sent:6 Jun 2013 01:52:50 +0000
To:'Info@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot
Dear Mr. Ihnot,

Thank you for your email below. We will review your Addendum 1 attachment and respond accordingly.

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: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Monday, June 03, 2013 9:41 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Dear Jill Constantine,

Thank you for the update. I appreciate that the WWC Quality Review Team has researched the issues related to WWC's reviews of Read Naturally, Inc. and Read Naturally® programs. I am grateful that you will be replacing the statement, "Read Naturally was found to have no discernible effects on reading fluency and comprehension." Read Naturally, Inc. has endured this inaccurate and damaging statement since 2006. I am excited that WWC will soon provide educators with an accurate description of Read Naturally® programs based on the Mesa, Christ, Heistad, Tucker and other studies.

Thank you for addressing many of my concerns in your emails. Your email from May 12, 2013 email does not address many of the issues I raised in Addendum 1 (attached) of my email to you on March 5. I still agree with Hancock, Hasbrouck (Denton Study), Kemp and Chenault that their studies should not be used as reviews of Read Naturally® programs as the studies were not primary analyses of the effect of an intervention. Addendum 1 spells out additional implementation variations. I have attached Addendum 1 to this email and hope you will review it.

However, the vast majority of educators will make judgments based on WWC's headline description. Thus, I am extremely happy that WWC will soon replace its inaccurate statement about Read Naturally, Inc. and Read Naturally® programs. I look forward to the revised report and updated statement.

Sincerely,

Tom Ihnot
CEO
Read Naturally, Inc.

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121

www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: WhatWorks [mailto:What.Works@icfi.com]
Sent: Thursday, May 30, 2013 2:42 PM
To: info@readnaturally.com
Cc: NSeftor@mathematica-mpr.com
Subject: What Works Clearinghouse (QR2010013)

Dear Mr. Ihnot,

The WWC is working on the update to the Beginning Reading Intervention Report on Read Naturally®, including revisiting the program description and program details to clarify the differences between Read Naturally, Inc., the Read Naturally strategy, and the four Read Naturally® products (ME, SE, Encore and Live). You will receive an embargoed copy of the report 24 hours before publication. Once released, the old report will be replaced on the web and the searches will report the findings in the updated report.

Thank you for continuing to bring to our attention your concerns about the studies reviewed by the WWC for the Read Naturally® intervention reports. We have previously responded to your concerns about the NCII reviews and the various studies (Mesa, Hancock, Kemp, Christ, Heistad, Denton, Kemp and Chenault). Please see our correspondence to you in February 2013 and May 2013, attached. If you have any different concerns, please let us know.

What Works Clearinghouse

WhatWorks

From:WhatWorks
Sent:17 Jun 2013 20:12:00 +0000
To:'Info@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot
Dear Mr. Ihnot,

Thank you for your email on June 3, 2013. We have reviewed Addendum 1, and as we have previously stated, our review concluded that the WWC followed protocol in determining whether the studies should be included in the intervention report. If you have any other different questions or concerns, please let us know.

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: Wednesday, June 05, 2013 8:53 PM
To: 'Info@readnaturally.com'
Subject: RE: Read Naturally - Tom Ihnot

Dear Mr. Ihnot,

Thank you for your email below. We will review your Addendum 1 attachment and respond accordingly.

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: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Monday, June 03, 2013 9:41 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Dear Jill Constantine,

Thank you for the update. I appreciate that the WWC Quality Review Team has researched the issues related to WWC's reviews of Read Naturally, Inc. and Read Naturally® programs. I am grateful that you will be replacing the statement, "Read Naturally was found to have no discernible effects on reading fluency and comprehension." Read Naturally, Inc. has endured this inaccurate and damaging statement since 2006. I am excited that WWC will soon provide educators with an accurate description of Read Naturally® programs based on the Mesa, Christ, Heistad, Tucker and other studies.

Thank you for addressing many of my concerns in your emails. Your email from May 12, 2013 email does not address many of the issues I raised in Addendum 1 (attached) of my email to you on March 5. I still agree with Hancock, Hasbrouck (Denton Study), Kemp and Chenault that their studies should not be used as reviews of Read Naturally® programs as the studies were not primary analyses of the effect of an intervention. Addendum 1 spells out additional implementation variations. I have attached Addendum 1 to this email and hope you will review it.

However, the vast majority of educators will make judgments based on WWC's headline description. Thus, I am extremely happy that WWC will soon replace its inaccurate statement about Read Naturally, Inc. and Read Naturally® programs. I look forward to the revised report and updated statement.

Sincerely,

Tom Ihnot
CEO
Read Naturally, Inc.

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: WhatWorks [mailto:What.Works@icfi.com]
Sent: Thursday, May 30, 2013 2:42 PM
To: info@readnaturally.com
Cc: NSeftor@mathematica-mpr.com
Subject: What Works Clearinghouse (QR2010013)

Dear Mr. Ihnot,

The WWC is working on the update to the Beginning Reading Intervention Report on Read Naturally®, including revisiting the program description and program details to clarify the differences between Read Naturally, Inc., the Read Naturally strategy, and the four Read Naturally® products (ME, SE, Encore and Live). You will receive an embargoed copy of the report 24 hours before publication. Once released, the old report will be replaced on the web and the searches will report the findings in the updated report.

Thank you for continuing to bring to our attention your concerns about the studies reviewed by the WWC for the Read Naturally® intervention reports. We have previously responded to your concerns

about the NCII reviews and the various studies (Mesa, Hancock, Kemp, Christ, Heistad, Denton, Kemp and Chenault). Please see our correspondence to you in February 2013 and May 2013, attached. If you have any different concerns, please let us know.

What Works Clearinghouse

Karla Ramy

From:Karla Ramy
Sent:24 Jun 2013 15:24:09 +0000
To:What Works
Subject:Read Naturally - Tom Ihnot
Jill Constantine,

Thank you for your email of June 17, 2013. What Works Clearinghouse (WWC) has clearly stated it does not consider implementation, which I think is a serious mistake. However, I do not think you have responded to my concerns about WWC's premise for selecting studies to review.

You have stated that WWC screens studies based on "whether the study represents a primary analysis of the effect of an intervention." It is clear Hancock, Denton, Kemp and Chenault did not do a primary analysis of the effect of a Read Naturally© intervention. In each case, the study authors said they happened to use passages published by Read Naturally, Inc. These passages are just one component of a Read Naturally© intervention and do not produce an effect when used in the absence of the Read Naturally© strategy.

Hancock's primary analysis was evaluating the trajectory of second grade fluency development, not the Read Naturally© strategy. Denton's choice of assessment shows her primary analysis was Read Well, not the Read Naturally© strategy. Kemp's selection of students above the 50th percentile in reading shows her primary analysis was Sustained Silent Reading, not the Read Naturally© strategy. Chenault's focus on writing shows her primary analysis was Pay Attention, not the Read Naturally© strategy. The authors of these studies have each stated clearly that their studies were not a primary analysis of the effect of a Read Naturally© intervention. Consequently, these studies must be dropped from WWC's review.

WWC lists deviations in the review summary that it discerns from the author's printed study. Not comparing what an author does to what is prescribed by a program causes WWC reviews to be inaccurate. It causes WWC to violate its own policy and review studies that are not a primary analysis of the effect of an intervention, such as the Hancock, Denton, Kemp, and Chenault studies.

Furthermore, it makes little difference if WWC lists implementation deviations since most educators will judge a program based on the highly visible summary statement. In Read Naturally, Inc.'s case, this summary statement, "Read Naturally was found to have no discernible effects on fluency and reading comprehension," unfairly convinces educators that our programs have no effect. This statement is based on studies that should not be included in WWC's review. In addition, it is inaccurate for the following reasons:

1. WWC recently posted a review of the Heistad Four School Study where WWC states that a Read Naturally, Inc. product showed a "statistically significant positive effect." The "no discernible effect" statement contradicts this review.
2. NCII is also under the United States Department of Education, and its reviews clearly state that the Read Naturally© strategy has positive effects. The WWC statement needs to be updated to reflect the evolving process of reviews.
3. In the original WWC Beginning Reading review, the Mesa study was inexplicably grouped with the Hancock study. The Mesa study should stand on its own. It clearly demonstrates the Read Naturally© strategy's significant positive effects in fluency and comprehension.
4. As you stated in your email, it is important to distinguish between Read Naturally, Inc. and the products used in the studies. Simply saying "Read Naturally was found…" has no benefit to

educators as they are looking for specific products, and our company offers dozens of solutions targeting different populations.

5. In both the Hancock and Kemp studies, the treatment group had fluency gains greater than average growth based on oral reading fluency norms. It is therefore inaccurate to state that the Read Naturally, Inc. program used had no discernible effect. However, these studies should not be used at all because they were not primary analyses of the effect of a Read Naturally© intervention.

6. When the Beginning Reading intervention report is updated, the Christ Study, Heistad Study (3rd grade), Mesa Study and any other control group case studies you choose to use will clearly demonstrate that the Read Naturally © strategy has discernible effects on fluency, rate, accuracy and comprehension. The current statement provides educators with an inconsistent message from WWC.

7. Educators typically do not have the time to analyze all the details in the intervention report on implementation. They see the headline. Now that WWC and NCII have posted new reviews on Read Naturally, Inc. products, it is important for educators to have the most up-to-date headline information.

WWC's written response system has been very frustrating. I appreciate that recently there has been an increased attempt to understand issues I have raised, but this correspondence has still been inadequate to fully discuss and resolve the problems associated with the current WWC reviews of Read Naturally © programs. Do I have another level of appeal?

I look forward to your response.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Read Naturally® CELEBRATING 20 YEARS

WhatWorks

From:WhatWorks
Sent:25 Jun 2013 15:27:49 +0000
To:'kramy@readnaturally.com';'Info@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot
Dear Mr. Ihnot,

Thank you for your email below. The quality review team is reviewing your most recent concerns and will respond.

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: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Monday, June 24, 2013 11:24 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Jill Constantine,

Thank you for your email of June 17, 2013. What Works Clearinghouse (WWC) has clearly stated it does not consider implementation, which I think is a serious mistake. However, I do not think you have responded to my concerns about WWC's premise for selecting studies to review.

You have stated that WWC screens studies based on "whether the study represents a primary analysis of the effect of an intervention." It is clear Hancock, Denton, Kemp and Chenault did not do a primary analysis of the effect of a Read Naturally© intervention. In each case, the study authors said they happened to use passages published by Read Naturally, Inc. These passages are just one component of a Read Naturally© intervention and do not produce an effect when used in the absence of the Read Naturally© strategy.

Hancock's primary analysis was evaluating the trajectory of second grade fluency development, not the Read Naturally© strategy. Denton's choice of assessment shows her primary analysis was Read Well, not the Read Naturally© strategy. Kemp's selection of students above the 50th percentile in reading shows her primary analysis was Sustained Silent Reading, not the Read Naturally© strategy. Chenault's focus on writing shows her primary analysis was Pay Attention, not the Read Naturally© strategy. The authors of these studies have each stated clearly that their studies were not a primary analysis of the effect of a Read Naturally© intervention. Consequently, these studies must be dropped from WWC's review.

WWC lists deviations in the review summary that it discerns from the author's printed study. Not comparing what an author does to what is prescribed by a program causes WWC reviews to be inaccurate. It causes WWC to violate its own policy and review studies that are not a primary analysis of the effect of an intervention, such as the Hancock, Denton, Kemp, and Chenault studies.

Furthermore, it makes little difference if WWC lists implementation deviations since most educators will judge a program based on the highly visible summary statement. In Read Naturally, Inc.'s case, this

summary statement, "Read Naturally was found to have no discernible effects on fluency and reading comprehension," unfairly convinces educators that our programs have no effect. This statement is based on studies that should not be included in WWC's review. In addition, it is inaccurate for the following reasons:

1. WWC recently posted a review of the Heistad Four School Study where WWC states that a Read Naturally, Inc. product showed a "statistically significant positive effect." The "no discernible effect" statement contradicts this review.
2. NCII is also under the United States Department of Education, and its reviews clearly state that the Read Naturally© strategy has positive effects. The WWC statement needs to be updated to reflect the evolving process of reviews.
3. In the original WWC Beginning Reading review, the Mesa study was inexplicably grouped with the Hancock study. The Mesa study should stand on its own. It clearly demonstrates the Read Naturally© strategy's significant positive effects in fluency and comprehension.
4. As you stated in your email, it is important to distinguish between Read Naturally, Inc. and the products used in the studies. Simply saying "Read Naturally was found…" has no benefit to educators as they are looking for specific products, and our company offers dozens of solutions targeting different populations.
5. In both the Hancock and Kemp studies, the treatment group had fluency gains greater than average growth based on oral reading fluency norms. It is therefore inaccurate to state that the Read Naturally, Inc. program used had no discernible effect. However, these studies should not be used at all because they were not primary analyses of the effect of a Read Naturally© intervention.
6. When the Beginning Reading intervention report is updated, the Christ Study, Heistad Study (3rd grade), Mesa Study and any other control group case studies you choose to use will clearly demonstrate that the Read Naturally© strategy has discernible effects on fluency, rate, accuracy and comprehension. The current statement provides educators with an inconsistent message from WWC.
7. Educators typically do not have the time to analyze all the details in the intervention report on implementation. They see the headline. Now that WWC and NCII have posted new reviews on Read Naturally, Inc. products, it is important for educators to have the most up-to-date headline information.

WWC's written response system has been very frustrating. I appreciate that recently there has been an increased attempt to understand issues I have raised, but this correspondence has still been inadequate to fully discuss and resolve the problems associated with the current WWC reviews of Read Naturally© programs. Do I have another level of appeal?

I look forward to your response.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190

St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

info@readnaturally.com

From:info@readnaturally.com
Sent:8 Jul 2013 17:16:15 +0000
To:What Works
Subject:RE: WWC Beginning Reading Review of Read Naturally
Jill,

I would like the chance to respond to your posting of the new Beginning Reading reviews. A one day notice is not enough time to evaluate the latest posting.

I request that the posting be delayed one week.

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: What Works [mailto:whatworks@mathematica-mpr.com]
Sent: Monday, July 08, 2013 9:01 AM
To: info@readnaturally.com
Subject: WWC Beginning Reading Review of Read Naturally

Dear Mr. Ihnot,

The attached letter is to notify you that the What Works Clearinghouse (WWC) has completed the review of the research on Read Naturally and determined that this intervention is eligible for an intervention report according to the Beginning Reading review protocol. We have also attached a courtesy copy of the report which will be posted on the WWC website on July 9, 2013. As a reminder, this report is covered by the embargo agreement signed by you on February 1, 2012, requiring you not to copy, distribute, or discuss the report with members of the public outside your organization, prior to release of the report by the Institute of Education Sciences.

Sincerely,
Jill Constantine
Director
What Works Clearinghouse at Mathematica Policy Research

WhatWorks

From:WhatWorks
Sent:9 Jul 2013 16:22:01 +0000
To:'Info@readnaturally.com'
Subject:RE: WWC Beginning Reading Review of Read Naturally
Dear Mr. Ihnot,

Thank you for your email to the What Works Clearinghouse (WWC) about the Beginning Reading review of Read Naturally. Our email of July 8, 2013 was an advance courtesy copy of the report. In keeping with WWC policy, we will release the report as scheduled. If you have specific comments or questions, please submit them in writing.

The quality review team is actively reviewing your previous correspondence and preparing a response to those concerns.

Thank you,

What Works Clearinghouse

From: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Monday, July 08, 2013 1:16 PM
To: What Works
Subject: RE: WWC Beginning Reading Review of Read Naturally

Jill,

I would like the chance to respond to your posting of the new Beginning Reading reviews. A one day notice is not enough time to evaluate the latest posting.

I request that the posting be delayed one week.

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: What Works [mailto:whatworks@mathematica-mpr.com]
Sent: Monday, July 08, 2013 9:01 AM
To: info@readnaturally.com
Subject: WWC Beginning Reading Review of Read Naturally

Dear Mr. Ihnot,

The attached letter is to notify you that the What Works Clearinghouse (WWC) has completed the review of the research on Read Naturally and determined that this intervention is eligible for an intervention report according to the Beginning Reading review protocol. We have also attached a courtesy copy of the report which will be posted on the WWC website on July 9, 2013. As a reminder, this report is covered by the embargo agreement signed by you on February 1, 2012, requiring you not to copy, distribute, or discuss the report with members of the public outside your organization, prior to release of the report by the Institute of Education Sciences.

Sincerely,
Jill Constantine
Director
What Works Clearinghouse at Mathematica Policy Research

From: info@readnaturally.com
Sent: Wednesday, July 17, 2013 3:16 PM
To: What Works
Subject: Tom Ihnot - Read Naturally

Jill Constantine,

The purpose of this letter is to request a minor change to the Effectiveness Summary paragraph for Read Naturally® programs in the new report summary for the Beginning Reading Evidence Review.
Specifically, I am requesting that WWC change the order in which the findings are listed.

I would appreciate if you change the Read Naturally® Effectiveness Summary paragraph to say:

Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

The content of the above paragraph is the same as the paragraph currently displayed on the WWC website; it just changes the order of the findings. The reasons for this change are as follows:

- Although Read Naturally, Inc. publishes a Phonics Series that specifically addresses the skills described in Alphabetics, the studies reviewed by What Works Clearinghouse used Read Naturally®'s Sequenced Series, which emphasizes general reading achievement and fluency. The Read Naturally® Sequenced Series does not explicitly address alphabetics, and comprehension assessments are not sensitive enough to pick up comprehension gains in short studies.
Consequently, it is not a useful summary of the Read Naturally® Sequenced Series to start by listing alphabetics and comprehension. The new report summary will be more helpful to educators if the paragraph lists positive effects (general reading achievement), mixed effects (reading fluency), and then no discernible effects (alphabetics and comprehension) to reflect the areas of reading the Read Naturally® Sequenced Series targets.
- Listing general reading achievement first, followed by reading fluency, followed by alphabetics and comprehension matches the order of the items in the Intervention Effectiveness Rating Chart that displays on the results page.
- Changing the order to list the positive findings first will allow Google and Bing search engine

users to immediately see Read Naturally®'s positive effects as they look for information about
our programs. Since 2007, when searching for Read Naturally® using Google and Bing search
engines, educators have first seen "Read Naturally was found to have no discernible effects…,"
which has been highly damaging to our company.
- WWC has reported statistically significant positive effects in fluency, accuracy, and general
reading achievement for Read Naturally® programs per the Christ study and Heistad studies.
Reporting the positive effects first in the Beginning Reading section would be consistent with
these findings, providing educators with a consistent message.

We know that What Works Clearinghouse is an important resource for educators, so we are eager for the
content to provide helpful, positive information to educators seeking information about Read Naturally®
programs.

I hope this change can be made as soon as possible.

Thank you,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:18 Jul 2013 22:28:05 +0000
To:'Info@readnaturally.com'
Subject:RE: Tom Ihnot - Read Naturally WWCPC 2304
Dear Mr. Ihnot,

The quality review team is reviewing your concerns and will send a response to you shortly.

What Works Clearinghouse

From: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Wednesday, July 17, 2013 3:16 PM
To: What Works
Subject: Tom Ihnot - Read Naturally

Jill Constantine,

The purpose of this letter is to request a minor change to the Effectiveness Summary paragraph for Read Naturally® programs in the new report summary for the Beginning Reading Evidence Review. Specifically, I am requesting that WWC change the order in which the findings are listed.

I would appreciate if you change the Read Naturally® Effectiveness Summary paragraph to say:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

The content of the above paragraph is the same as the paragraph currently displayed on the WWC website; it just changes the order of the findings. The reasons for this change are as follows:

- Although Read Naturally, Inc. publishes a Phonics Series that specifically addresses the skills described in *Alphabetics*, the studies reviewed by What Works Clearinghouse used Read Naturally®'s Sequenced Series, which emphasizes general reading achievement and fluency. The Read Naturally® Sequenced Series does not explicitly address alphabetics, and comprehension assessments are not sensitive enough to pick up comprehension gains in short studies. Consequently, it is not a useful summary of the Read Naturally® Sequenced Series to start by listing alphabetics and comprehension. The new report summary will be more helpful to educators if the paragraph lists *positive effects* (general reading achievement), *mixed effects* (reading fluency), and then *no discernible effects* (alphabetics and comprehension) to reflect the areas of reading the Read Naturally® Sequenced Series targets.
- Listing general reading achievement first, followed by reading fluency, followed by alphabetics and comprehension matches the order of the items in the Intervention Effectiveness Rating Chart that displays on the results page.
- Changing the order to list the positive findings first will allow Google and Bing search engine users to immediately see Read Naturally®'s positive effects as they look for information about our programs. Since 2007, when searching for Read Naturally® using Google and Bing search engines, educators have first seen "Read Naturally was found to have no discernible effects…," which has been highly damaging to our company.
- WWC has reported statistically significant positive effects in fluency, accuracy, and general reading achievement for Read Naturally® programs per the Christ study and Heistad studies.

Reporting the positive effects first in the Beginning Reading section would be consistent with these findings, providing educators with a consistent message.

We know that What Works Clearinghouse is an important resource for educators, so we are eager for the content to provide helpful, positive information to educators seeking information about Read Naturally® programs.

I hope this change can be made as soon as possible.

Thank you,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Karla Ramy

From:Karla Ramy
Sent:5 Aug 2013 19:33:53 +0000
To:What Works
Subject:Tom Ihnot - Read Naturally
Jill Constantine,

I am hopeful about a positive response to my 7/17/13 request (see below). What Works Clearinghouse (WWC) 7/18/13 response indicated that the quality review team would send their response shortly.

Also, I would like to know why the fluency results in the Heistad study were not used in the fluency evaluation.

Thank you,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: info@readnaturally.com
Sent: Wednesday, July 17, 2013 2:16 PM
To: 'What Works'
Subject: Tom Ihnot - Read Naturally

Jill Constantine,

The purpose of this letter is to request a minor change to the Effectiveness Summary paragraph for Read Naturally® programs in the new report summary for the Beginning Reading Evidence Review. Specifically, I am requesting that WWC change the order in which the findings are listed.

I would appreciate if you change the Read Naturally® Effectiveness Summary paragraph to say:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

The content of the above paragraph is the same as the paragraph currently displayed on the WWC website; it just changes the order of the findings. The reasons for this change are as follows:

- Although Read Naturally, Inc. publishes a Phonics Series that specifically addresses the skills described in *Alphabetics*, the studies reviewed by What Works Clearinghouse used Read Naturally®'s Sequenced Series, which emphasizes general reading achievement and fluency. The Read Naturally® Sequenced Series does not explicitly address alphabetics, and comprehension assessments are not sensitive enough to pick up comprehension gains in short studies. Consequently, it is not a useful summary of the Read Naturally® Sequenced Series to start by listing alphabetics and comprehension. The new report summary will be more helpful to educators if the paragraph lists *positive effects* (general reading achievement), *mixed effects* (reading fluency), and then *no discernible effects* (alphabetics and comprehension) to reflect the areas of reading the Read Naturally® Sequenced Series targets.
- Listing general reading achievement first, followed by reading fluency, followed by alphabetics and comprehension matches the order of the items in the Intervention Effectiveness Rating Chart that displays on the results page.
- Changing the order to list the positive findings first will allow Google and Bing search engine users to immediately see Read Naturally®'s positive effects as they look for information about our programs. Since 2007, when searching for Read Naturally® using Google and Bing search engines, educators have first seen "Read Naturally was found to have no discernible effects…," which has been highly damaging to our company.
- WWC has reported statistically significant positive effects in fluency, accuracy, and general reading achievement for Read Naturally® programs per the Christ study and Heistad studies. Reporting the positive effects first in the Beginning Reading section would be consistent with these findings, providing educators with a consistent message.

We know that What Works Clearinghouse is an important resource for educators, so we are eager for the content to provide helpful, positive information to educators seeking information about Read Naturally® programs.

I hope this change can be made as soon as possible.

Thank you,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:7 Aug 2013 00:40:01 +0000
To:'kramy@readnaturally.com';'info@readnaturally.com'
Subject:RE: Tom Ihnot - Read Naturally
Dear Mr. Ihnot,

We expect to have the Quality Review Team response to you within the next several days.

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Monday, August 05, 2013 3:34 PM
To: What Works
Subject: Tom Ihnot - Read Naturally

Jill Constantine,

I am hopeful about a positive response to my 7/17/13 request (see below). What Works Clearinghouse (WWC) 7/18/13 response indicated that the quality review team would send their response shortly.

Also, I would like to know why the fluency results in the Heistad study were not used in the fluency evaluation.

Thank you,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: info@readnaturally.com
Sent: Wednesday, July 17, 2013 2:16 PM

To: 'What Works'
Subject: Tom Ihnot - Read Naturally

Jill Constantine,

The purpose of this letter is to request a minor change to the Effectiveness Summary paragraph for Read Naturally® programs in the new report summary for the Beginning Reading Evidence Review. Specifically, I am requesting that WWC change the order in which the findings are listed.

I would appreciate if you change the Read Naturally® Effectiveness Summary paragraph to say:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

The content of the above paragraph is the same as the paragraph currently displayed on the WWC website; it just changes the order of the findings. The reasons for this change are as follows:

- Although Read Naturally, Inc. publishes a Phonics Series that specifically addresses the skills described in *Alphabetics*, the studies reviewed by What Works Clearinghouse used Read Naturally®'s Sequenced Series, which emphasizes general reading achievement and fluency. The Read Naturally® Sequenced Series does not explicitly address alphabetics, and comprehension assessments are not sensitive enough to pick up comprehension gains in short studies. Consequently, it is not a useful summary of the Read Naturally® Sequenced Series to start by listing alphabetics and comprehension. The new report summary will be more helpful to educators if the paragraph lists *positive effects* (general reading achievement), *mixed effects* (reading fluency), and then *no discernible effects* (alphabetics and comprehension) to reflect the areas of reading the Read Naturally® Sequenced Series targets.
- Listing general reading achievement first, followed by reading fluency, followed by alphabetics and comprehension matches the order of the items in the Intervention Effectiveness Rating Chart that displays on the results page.
- Changing the order to list the positive findings first will allow Google and Bing search engine users to immediately see Read Naturally®'s positive effects as they look for information about our programs. Since 2007, when searching for Read Naturally® using Google and Bing search engines, educators have first seen "Read Naturally was found to have no discernible effects…," which has been highly damaging to our company.
- WWC has reported statistically significant positive effects in fluency, accuracy, and general reading achievement for Read Naturally® programs per the Christ study and Heistad studies. Reporting the positive effects first in the Beginning Reading section would be consistent with these findings, providing educators with a consistent message.

We know that What Works Clearinghouse is an important resource for educators, so we are eager for the content to provide helpful, positive information to educators seeking information about Read Naturally® programs.

I hope this change can be made as soon as possible.

Thank you,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com

www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Karla Ramy

From:Karla Ramy
Sent:7 Aug 2013 20:19:14 +0000
To:What Works
Subject:Tom Ihnot
Jill Constantine,

I would like to clarify my question concerning the fluency results in the Heistad study that I emailed on 8/5/13. To be more precise, did What Works Clearinghouse (WWC) exclude the Heistad fluency results because he stated in his analysis that he used Read Naturally's Reading Fluency Monitor passages to asses for fluency gains?

Thank you,
Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:8 Aug 2013 01:54:34 +0000
To:'kramy@readnaturally.com'
Subject:RE: Tom Ihnot
Dear Mr. Ihnot,

We passed this clarification on to the quality review team and they will address this question when finalizing a response to your concerns.

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Wednesday, August 07, 2013 4:19 PM
To: What Works
Subject: Tom Ihnot

Jill Constantine,

I would like to clarify my question concerning the fluency results in the Heistad study that I emailed on 8/5/13. To be more precise, did What Works Clearinghouse (WWC) exclude the Heistad fluency results because he stated in his analysis that he used Read Naturally's Reading Fluency Monitor passages to asses for fluency gains?

Thank you,
Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

School Psychology Review,
2010, Volume 39, No. 3, pp. 447–462

Curriculum-Based Measurement of Oral Reading: An Evaluation of Growth Rates and Seasonal Effects Among Students Served in General and Special Education

Theodore J. Christ
University of Minnesota

Benjamin Silberglitt
Technology and Information Educational Services

Seungsoo Yeo
Korean Educational Development Institute

Damien Cormier
University of Minnesota

Abstract. Curriculum-based measurement of oral reading (CBM-R) is often used to benchmark growth in the fall, winter, and spring. CBM-R is also used to set goals and monitor student progress between benchmarking occasions. The results of previous research establish an expectation that weekly growth on CBM-R tasks is consistently linear throughout the academic year. The patterns of CBM-R growth were examined for a large sample of students (N = 3808) from both general education and special education populations in second to sixth grades. Results support four general conclusions: (a) annual growth is more substantial within the general education population; (b) growth is more substantial in earlier elementary grades; (c) more growth occurs in the fall than the spring season (i.e., seasonal effect), especially within the early primary general education population; and (d) the seasonal effect is less pronounced within the special education population. Estimates of growth within and across seasons are presented and implications are discussed.

Curriculum-based measurement (CBM) is used to index annual student growth across the primary grades. Procedures and measurement metrics are developed for mathematics, spelling, written expression, and reading. CBM oral reading (CBM-R) rate is the most researched and well established of those available procedures (Wayman, Wallace, Wiley, Ticha, & Espin, 2007). CBM-R is used as an index of reading rate and fluency, which are identified as critical skills to target for instruction and intervention within the early stages of

Correspondence regarding this article should be addressed to Theodore J. Christ, 56 East River Road, Minneapolis, MN 55455; E-mail: tchrist@umn.edu

Copyright 2010 by the National Association of School Psychologists, ISSN 0279-6015

reading development (National Reading Panel, 2000; National Research Council, 1998). CBM-R outcome also functions as a robust indicator of overall reading development throughout the primary grades (Wayman et al., 2007).

CBM-R progress monitoring data are collected and used to evaluate instructional effects, determine when to modify the instruction, and evaluate response to intervention (Deno, 1985, 1986). Teachers who use progress monitoring data are likely to use more specific and measurable goals, rely more substantially on data to guide instruction, and modify instruction more frequently (Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, & Stecker, 1989). Progress monitoring applications of CBM-R have the potential to confer the most substantial benefit to students. Progress monitoring can be conceptualized as either interim benchmark/screening assessments when data are collected three to four times per year, or time series continuous assessments when data are collected daily or weekly.

Both interim and time series CBM-R data are plotted graphically and evaluated against goals and goal lines. Fuchs and Shinn (1989) recommended that the rate of expected growth should derive from local normative performance so that individual student achievement can be compared against the local normative performance of grade-level peers. A goal line is established by connecting the CBM-R data point for the observed level of performance from Week 1 (initial performance) to the expected/typical level of performance at Week 36 (ending performance). Once graphed, the goal line provides a graphic trajectory of expected growth for the academic year. Ongoing time series data are then plotted on the graph and evaluated against the goal line. For example, if a student's initial level of CBM-R performance in September was 20 words read correctly per minute (WRCM) and in May that student was expected to read 60 WRCM, then the expected annual growth is 40 WRCM across 8 months, or 32 weeks. That translates to ~1.25 WRCM per week of growth (1.25 WRCM = 40 WRCM/32 weeks). The goal line and trajectory of expected growth is based on the assumption of a monotonic and linear trend.

Monotonic Linear Growth

Standards of expected performance within and across the year provide the foundation to establish instructional goals and goal lines. There are two influential studies that established standards of weekly CBM-R growth (Deno, Fuchs, Marston, & Shin, 2001; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993), which have application in both research and practice. Fuchs et al. (1993) analyzed a CBM-R data set comprised of 103 students in general education and 14 students in special education. Each student was assessed weekly for an entire academic year using CBM-R procedures. Although some students evidenced a negatively accelerating annual growth, the researchers concluded that annual growth is, on average, monotonic and linear. They also observed a negatively accelerating trend across grades. That is, the rate of CBM-R growth was positive in all grades, but the rate of growth declined in each successive grade. The resulting analysis yielded grade-specific estimates of standard and ambitious rates of growth—and *assumed* a pattern of monotonic linear growth. Respectively, the growth, determined by the number of WRCM, were 1.5 and 2.0 in first and second grade, 1.0 and 1.5 in third, 0.85 and 1.1 in fourth, 0.5 and 0.8 in fifth, and 0.3 and 0.63 in sixth.

Deno et al. (2001) analyzed a large data set comprised of 2675 students in general education and 324 in special education classes. The student sample was from school districts across the country. The researchers relied on the *assumption* of monotonic linear annual growth. Ordinary least squares regression was used to calculate weekly growth. The study yielded useful estimates of observed and expected growth across grades and populations. The authors estimated that beginning readers in the general education population can be expected to improve 2 WRCM/week until they achieve 30 WRCM; thereafter, students in the general education population can be expect to improve at least 1 WRCM/week. Researchers observed that students in the special education

population demonstrated substantially less growth, which approximated 0.5–0.8 WRCM/week, unless they are provided with robust evidence-based instruction, which might improve the observed rate of growth to 1.39 WRCM/week.

Prior studies provide useful estimates of typical, ambitious, and expected rates of CBM-R growth; however, those estimates depend on the assumption that growth is monotonic and linear across the academic year. That assumption should be examined more closely to determine its veracity. The prior studies also rely on analyses that neglect the inherent nesting of students within classrooms and schools, which establish potential bias in the linear and quadratic functions derived.

Improved Modeling of Annual and Seasonal Growth

Research is necessary to evaluate alternate models of annual growth and test the predominant assumption that annual growth is monotonic and linear. This study will evaluate both linear (constant) and piece wise (nonconstant) models of growth from fall to winter to spring; moreover, this study will employ the superior approach of linear mixed model (LMM) rather than that of ordinary least squares used in prior research (Deno et al., 2001; Fuchs et al., 1993).

Ordinary least squares only estimates the fixed effects, which assumes that all cases have identical parameter estimates for intercept and growth. This is a limitation when modeling CBM-R growth because it is likely to be an erroneous assumption in most cases. The results of published research provide evidence for substantial magnitudes of variance in both intercept and growth among students in a sample (Shin, Deno, & Espin, 2000; Stage, 2001). It is likely that the intercept and growth rates should be treated as random effects in many instances. LMM can be used to estimate both the fixed effects and the random effects, which are the intercepts and growth rates for individual student cases in the sample. LMM takes into account the variances within and across individuals in the sample. Finally, ordinary least squares requires that missing data are removed casewise whereas LMM can handle missing data, which is likely to yield less biased estimates of intercept and slopes while enhancing power (Duncan, Duncan, & Strycker, 2006). This approach facilitates analysis of both seasonal effects, as pursued by two previous studies (Ardoin & Christ, 2008; Graney, Missall, Martinez, & Bergstrom, 2009), and compares those effects to that of linear growth as assumed in other studies (Deno et al., 2001; Fuchs et al., 1993).

The calendar year was divided into three seasons for the purpose of this study. The seasons were defined by those periods between standard tri-annual assessment occasions: fall, winter and spring. The duration between fall and winter assessments defines the fall season, that between winter and spring defines the spring season, and that between spring and fall defines the summer season. If growth is consistent and linear throughout the calendar year, then there is no *seasonal effect*. Conversely, if growth is not consistent and linear throughout the calendar year, then there is a seasonal effect whereby there is more growth observed for some seasons than for others.

The summer season is the most distinct for the majority of students because they are not exposed to formal instruction. The rate of achievement is likely to decline in the summer months (Kim, 2004) unless formal instruction is provided through summer school programs (Stage, 2001). A seasonal effect in the summer is expected because the conditions for learning are substantially different from that which occurs in the spring and fall seasons. The effect, however, does not seem specific to those differences between summer and the academic year. Ardoin and Christ (2008) observed statistically significant differences in CBM-R growth between the fall and spring seasons. That is, among a sample of second-grade students there was more growth in the fall than in the spring season. That result replicated across three combinations of CBM-R passages sampled from the Dynamic Indicators of Basic Early Literacy Skills. A subsequent study aimed to replicate and extend those findings

with a sample of third- through fifth-grade students across a variety of CBM curriculum domains (Graney et al., 2009). However, Graney et al. (2009) concluded that there was more CBM-R growth in the spring season than in the fall season, failing to replicate the work of Ardoin and Christ. Graney et al. speculated that the failure to replicate findings across studies might relate to the specification of data collection schedules, sample characteristics, issues of instrumentation, or the distribution of instructional intensity across the academic year. Although it is difficult to make a single attribution, it should be noted that the student sample used by Graney et al. derived from a school–university partnership program that included features to promote data use and evidence-based interventions. Indeed, that program might influence the pattern of student growth; however, additional research and replications are necessary to help determine whether specific patterns for seasonal effects should be expected.

There are limitations to the analytic procedures used in the two prior studies. Both generated estimates of growth from difference scores across seasons. That is, growth was estimated as the difference between performance at Time 1 and Time 2, which was then divided by the number of intervening weeks to yield an estimate of weekly growth. Although this method is intuitive, and it was useful for the preliminary work to examine seasonal effects, the method of difference scores to estimate change has long been criticized as an unreliable method of analysis (Cronbach & Furby, 1970) and may have contributed to the inconsistent findings. The analytic procedures used in this study improve on those of prior studies that examined seasonal effects and modeled annual growth. The analytic procedures used in this study will compare the hypothesized seasonal effect directly with that of linear growth.

Purpose

The purposes of this study were to (a) estimate the rate of growth within and across seasons for general education and special education populations; (b) evaluate whether the rate of CBM-R growth is likely to be consistent (i.e., linear model) or inconsistent (i.e., piecewise model) throughout the academic year; and (c) test the hypothesis that the seasonal effect is typically characterized by more robust growth in the fall than for the spring season. That hypothesis is counter to the assumptions inherent to the linear models as applied in many previous studies and it is counter to the findings of Graney et al. (2009).

Method

Participants and Setting

A total of 4824 students (52% male, 48% female) in the second through sixth grades were included in the sample. The ethnicity breakdown of the sample was 94% White, not Hispanic; 2% Native American; 2% Asian or Pacific Islander; 1% Hispanic; and 1% Black, not Hispanic. Participants were enrolled in seven elementary schools, located within five school districts in the rural and suburban Midwest. The sample consisted of 31% students receiving free or reduced-price lunch, and 8% students receiving special education services.

As part of an ongoing benchmark assessment plan and response to intervention implementation (see Bollman, Silberglitt, & Gibbons, 2007, for additional information), all students in participating schools were assessed using CBM-R every fall, winter, and spring of every year across the grade levels studied. All students for whom relevant data were collected participated in the study. Data were gathered between the years of 2001 and 2005 for each student in the sample. Data from across multiple years were available for many of the students in the sample; thus many students participated in the study at multiple grade levels.

Measures

The CBM-R measures were developed or selected to be curriculum-independent, grade-level appropriate, and of equivalent difficulty across probes within each grade level. Each probe contained approximately 250

words typed on an 8 by 10 piece of paper in a grade-appropriate font (range = 12–14 point with larger print for the second and third grades). Probes were selected from a standardized set of commercially available published passages (AIMSweb; Howe & Shinn, 2002). The technical manual for AIMSweb passages (Howe & Shinn, 2002) and peer-reviewed research (Wayman et al., 2007) provide evidence that test–retest reliability and alternate-form reliabilities approximate or exceed .90. Criterion-related validity coefficients typically approximate the .70–.90 range. The same set of CBM-R probes were used across all districts for all 4 years of the study.

Procedure

CBM-R data were collected in 4-month intervals in the fall, winter, and spring of each academic year. Standardized administration and scoring procedures were implemented for each administration, which was conducted by trained school-based personnel and high school honors students. These individuals were trained to administer three successive CBM-R probes and score WRCM using standardized procedures that have appeared in the published literature (Howe & Shinn, 2002; Shinn, 1988, 1989). Administration occurred outside of the classroom to minimize distractions, in a separate room or quiet hallway while the administrator and student were seated at desks or in chairs.

Training included a brief instructional session that lasted approximately 1 hr. Each training session was followed by an assessment of competency that included practice and evaluation of scoring procedures. No formal criterion for scoring was established until the final years of the study. In 2003–2004 and 2004–2005, each scorer was required to come within 2 WCPM of the correct score on three consecutive videotaped assessments. Although data were not maintained, few scorers required more than one opportunity to meet criterion. Finally, inter-rater reliability was not evaluated; however, CBM-R procedures have established high levels of inter-rater reliability that appear consistently within the published literature. A survey of 10 arbitrarily selected CBM-R studies that were published between 1995 and 2005 reported inter-rater reliabilities that were consistently above .95.

Analytical Plan

A LMM analysis was performed to investigate evidence of seasonal effects for CBM-R. The present LMM analysis considered two hypothesized growth models: linear and piecewise (Duncan et al., 2006; Flora, 2008). The linear model assumes a static growth rate across time points—meaning that regardless of change over time, the estimated growth rate from the linear model is identical to the sample's growth rate. In contrast, the piecewise model assumes that the developmental trajectory is not consistent across time, making it possible to provide two different linear trends. Thus, the piecewise model takes into account a possible nonlinear trajectory for CBM-R. Based on the three CBM-R data points used in this study, the piecewise model can be parameterized into two levels of analyses. First, Level 1 can be expressed as

$$y_{it} = \beta_{0i} + \beta_{1i}\lambda_{1t} + \beta_{2i}\lambda_{2t} + \varepsilon_{it} \quad (1)$$

where y_{it} is the outcome of CBM-R for subject i at time t, β_{0i} is the value of CBM-R for subject i in fall, β_{1i} is the first linear trend of subject i from fall to winter, λ_{1t} is the value of the linear trend at time t (Week 0 and Week 18), β_{2i} is the second linear trend of subject i from winter to spring, λ_{2t} is the value of the linear trend at time t (Week 18 and Week 36), and ε_{it} is unique measurement error for subject i at time t.

Level 2 can be described as

$$\beta_{0i} = \beta_0\zeta_{0i} \quad (2)$$

$$\beta_{1i} = \beta_1\zeta_{1i} \quad (3)$$

$$\beta_{2i} = \beta_2\zeta_{2i} \quad (4)$$

where β_0 represents the average initial status for all subjects, β_1 refers to the average of the first linear growth rate for all subjects, and β_2

refers to the average of the second slope trend for all subjects. ζ_{0i}, ζ_{1i}, and ζ_{2i} refer to the random deviation of the initial status, the first growth rate, and the second growth rate, respectively. As a result, the piecewise model presented earlier assumes that the second slope is statistically different from the first slope, indicating that there is a seasonal effect for CBM-R throughout the academic year.

The first step in this analysis was to compare the two models (linear and piecewise models) using fit indices, which included the likelihood ratio test (LRT), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The three fit indices are considered to be "goodness-of-fit" indices, where the smallest fit statistics are viewed as the best model. If results from these goodness-of-fit indices reveal that the linear model is better than the piecewise model, then the slope is the same from fall to spring and there is no evidence of a significant seasonal effect for CBM-R across the academic year. Conversely, if goodness-of-fit indices are better for the piecewise model than the linear model, then we conclude that the slopes are different from fall to winter and from winter to spring and there is evidence of a seasonal effect across the academic year. If the piecewise model is the better fit, then an additional model is needed to estimate the extent to which the first slope differs from the second slope.

The additional model, which is called *added growth model,* would be identical to the piecewise model described earlier, except for the supplementary time factor loading matrix (Flora, 2008). For example, in the piecewise model, the time factor loading can be described as

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 18 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (5)$$

where the first column is set at 1, meaning that the outcome of the fall CBM-R score (intercept) does not change across time. The second column refers to the first linear time metric indicating change in CBM-R score from fall to winter (Flora, 2008). Similarly, the third column represents change in CBM-R score from winter to fall.

For the added growth model, the time factor loading can be expressed as:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 36 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (6)$$

The matrix in Equation 6 is identical to the matrix in Equation 5, with the exception of the second column with 0, 18, and 36. The different time factor loadings for the additional model is applied to examine the extent to which linear change in CBM-R from fall to winter increases or decreases compared to that from winter to spring. In other words, the added growth model analysis makes it possible to contrast the first linear slope and the second linear slope. The results from the additional analysis can also be viewed as an effect size estimate between the first and second linear slope (Flora, 2008).

Data Preparation

Overall, the rates of missing data for fall, winter, and spring seasons were less than 4% for all grade levels. All LMM analyses were conducted using the SAS MIXED program with full maximum likelihood estimation that automatically conducts all adjustments for missing data. Monthly CBM data used in this study were nested within districts. Therefore, a three-level LMM was used to satisfy an assumption that data collected in this study might not be independent among school districts. An alpha level for all statistical tests in this study was set at .001.

Model assumptions were evaluated, which included a test of normality using Kolmogorov-Smirnove statistics, skew, and kurtosis, and homoscedasticity of residuals by plotting raw residuals at each level (Levels 1 and 2). No violations were detected. The Level 3 LMM analysis was used to test whether CBM data used in the study were nested within school districts. As a result, variations in the slopes and intercepts between school districts were not significantly different

Table 1
Descriptive Statistics for the Level and Change of CBM-R

	Seasonal Benchmark Levels			Weekly Change	
Population	Fall *M* (*SD*)	Winter *M* (*SD*)	Spring *M* (*SD*)	Fall Season (Fall–Winter) *M* (*SD*)	Spring Season (Winter–Spring) *M* (*SD*)
General education					
Second	57 (32)	88 (34)	107 (35)	1.71 (.73)	1.02 (.64)
Third	82 (37)	107 (38)	125 (40)	1.38 (.79)	0.97 (.75)
Fourth	106 (38)	128 (39)	141 (41)	1.19 (.72)	0.78 (.70)
Fifth	125 (40)	145 (41)	160 (42)	1.12 (.76)	0.83 (.75)
Sixth	142 (41)	158 (42)	171 (43)	0.89 (.71)	0.74 (.75)
Total	99 (47)	122 (45)	137 (46)	1.26 (.74)	0.87 (.72)
Special education					
Second	33 (28)	55 (35)	74 (38)	1.17 (.81)	1.08 (.63)
Third	52 (33)	72 (39)	88 (41)	1.09 (.81)	0.89 (.78)
Fourth	77 (40)	96 (43)	108 (45)	0.99 (.72)	0.67 (.65)
Fifth	89 (43)	106 (45)	118 (47)	0.90 (.76)	0.67 (.79)
Sixth	106 (47)	119 (50)	132 (51)	0.70 (.72)	0.73 (.69)
Total	72 (45)	89 (48)	104 (49)	0.97 (.77)	0.81 (.71)

Note. CBM-R = curriculum-based measurement of oral reading. Skew and kurtosis were within reasonable limits ($z \leq 2.00$) for all conditions. *N* values for second, third, fourth, fifth, and sixth grades were, respectively, 1790, 1885, 1913, 1942, and 1144 for general education; 111, 134, 170, 192, and 116 for special education.

from zero for all grade levels, meaning that both the slopes and intercepts at the district levels were identical. In addition, such a result supports evidence that the Level 2 LMM analysis introduced in the method section is appropriate for this study.

Results

Summaries of the means and standard deviations for CBM-R at each time point are presented in Table 1. As shown in Table 1, the average growth as measured with CBM-R was positive for all conditions across seasons (fall–winter, winter–spring), grade levels (second to sixth) and populations (general education, special education).

Comparison Between Linear Model and Piecewise Model

Results for the comparisons of linear and piecewise models are presented in Table 2. For students in general education, the piecewise model provided the best fit for tri-annual assessment data, as evidenced by the magnitudes of LRT, AIC, and BIC observed for the linear and piecewise models respectively. The χ^2 difference test between the two models also revealed that the piecewise model fit significantly better than the linear model.

For students in special education, on the basis of the LRT criterion, the piecewise model for all grade levels was viewed as the best-fitting model. Significant differences in LRTs between the two models, however, were found in the second through fourth grades. There were inconsistencies with respect to AIC and BIC criteria, which might indicate the linear model fit better in some cases within the sample of special education students. Although the results of AIC and BIC are noted, the LRT was the primary test/criterion used here.

Table 2
Deviation Between Linear and Piecewise Models for General Education Students and Special Education Students

	Linear Model			Piecewise Model			Differences (Linear-Piecewise Model)		
Grade	LRT	AIC	BIC	LRT	AIC	BIC	LRT	AIC	BIC
General education									
Second	46231.7	46239.7	46261.7	45542.6	45556.6	45595.0	689.1*	683.1	666.7
Third	49566.9	49574.9	49597.0	49357.7	49371.7	49410.5	209.2*	203.2	186.5
Fourth	49768.1	49776.1	49798.3	49520.1	49534.1	49573.0	248.0*	242.0	225.3
Fifth	50974.9	50982.9	51005.2	50866.6	50880.6	50919.6	108.3*	102.3	85.6
Sixth	29892.9	29900.9	29921.1	29873.6	29887.6	29922.9	19.3*	13.3	−1.8
Special education									
Second	2816.3	2824.3	2835.1	2802.6	2816.6	2835.6	13.7*	7.7	−0.5
Third	3501.8	3507.8	3516.5	3487.9	3495.9	3507.5	13.9*	11.9	9.0
Fourth	4394.9	4402.9	4415.4	4378.9	4392.9	4414.9	16.0*	10.0	0.5
Fifth	5090.6	5098.6	5111.6	5083.3	5097.3	5120.1	7.3	1.3	−8.5
Sixth	3048.2	3056.2	3067.2	3046.3	3060.3	3079.5	1.9	−4.1	−12.3

Note. LRT = likelihood ratio test; AIC = Akaike information criterion; BIC = Bayesian information criterion.
*$p < .001$.

Comparison Between the First Slope and Second Slope for General and Special Education

Table 3 presents estimates of both the first (fall–winter) and second (winter–spring) linear slope. In the cases of both students in general education and special education, all of the seasonal linear slopes were positive with magnitudes significantly different than zero ($p < .001$). That observation was consistent with the conclusion that reading achievement accelerated continuously throughout the academic year for all student samples. Although the mean rate of growth across seasons was consistently greater for students in general education as compared with students in special education, it was the first slopes that were most distinct. Further, the differences in the second slopes varied by grade level.

Additional analysis was conducted to examine whether the first linear slopes (fall) were significantly different from the second slopes (spring). Table 3 presents the differences between the two linear slopes along with the results of significance tests. Figure 1 depicts the slopes and model differences.

In the sample of general education students, statistically significant differences ($p < .001$) were observed between the first and second slopes in all grade-level conditions (Table 3). That is, the rate of growth for all grade levels dropped in the second part of the academic year. It is relevant to note that the magnitude of difference between the two slopes (seasonal effects) and standardized effect sizes consistently decreased as grade level increased. This might indicate that the seasonal effect in reading is larger in the lower primary grades than the upper primary grades. The power estimates for detecting seasonal effects for students in general education were very large for all grade levels (range = 0.96–1.00). This indicates a good chance of detecting a false null hypothesis (linear model) when the alternative model (piecewise model) is correct (Duncan et al., 2006).

In the sample of special education students, the average of the first slope (fall) was

Table 3
Comparison Between the First Slope and the Second Slope for General Education Students and Special Education Students

Grade	Intercept $\hat{\beta}_0$ (*SE*)	Fall/First Slope (Fall to Winter) $\hat{\beta}_1$ (*SE*)	Spring/ Second Slope (Winter to Spring) $\hat{\beta}_2$ (*SE*)	Difference Between Fall Slope and Spring Slopes $\hat{\beta}_2{}^a$ (*SE*)	Standardized Effect Sizes[b]	Power
General education						
Second	57.42* (.76)	1.71* (.02)	1.02* (.01)	0.69* (.02)	34.50	1.00
Third	82.31* (.85)	1.38* (.02)	0.97* (.02)	0.41* (.03)	13.60	1.00
Fourth	105.78* (.93)	1.19* (.02)	0.78* (.02)	0.41* (.12)	3.41	1.00
Fifth	124.43* (.89)	1.12* (.02)	0.83* (.02)	0.29* (.12)	2.41	1.00
Sixth	142.49* (1.21)	0.88* (.02)	0.74* (.02)	0.14* (.16)	0.87	0.96
Special education						
Second	34.02* (2.67)	1.17* (.08)	1.08* (.06)	0.09 (.10)	0.90	0.86
Third	50.85* (2.81)	1.09* (.07)	0.89* (.06)	0.20 (.10)	2.00	0.87
Fourth	75.50* (2.99)	0.99* (.05)	0.67* (.05)	0.32* (.08)	4.00	0.92
Fifth	86.28* (3.06)	0.90* (.05)	0.67* (.06)	0.23 (.09)	2.50	0.54
Sixth	106.98* (4.40)	0.69* (.07)	0.73* (.06)	−0.04 (.10)	0.40	0.11

Note. SE = standard error.
[a]The estimated coefficients of the second slope from an added growth model.
[b]Standardized effect sizes = $\hat{\beta}_2$/SE, which is a strategy developed by Satorra and Saris (1985), was used to calculate power.
*$p < .001$.

greater than the second slope (spring) in all grades except for sixth; however, those differences were not statistically significant except in the fourth-grade condition (Table 3). Moreover, the pattern of reduced seasonal effects for progressively higher grades did not generalize to the sample of special education students. Instead, the magnitude was greatest in fourth grade (difference in slopes = 0.32, $p < .001$), somewhat smaller in third (0.20, $p > .001$) and fifth (0.23, $p > .001$), and hardly perceptible in second (0.09, $p > .001$) and sixth grades (−0.04, $p < .001$). The range of standardized effect sizes and power was variable across grades. Based on the results, it appears that, in general, the observed slopes in the sample of special education students were relatively consistent across seasons—in all conditions but fourth grade.

Discussion

Many of the published estimates of CBM-R annual growth rates rely on the assumption that the pattern is consistent and linear throughout the academic year (Deno et al., 2001; Fuchs et al., 1993). Although there is evidence for seasonal effects within some studies (Ardoin & Christ, 2009; Fuchs et al., 1993; Graney et al., 2009), the phenomenon of seasonal effects are generally ignored within the professional literature. This study yielded (a) estimates for the magnitude CBM-R annual growth along with those within the fall and spring seasons, and estimates were derived for both general education and special education populations; (b) both linear and piecewise growth models; and (c) hypothesis tests to examine whether fall growth is typically more substantial than spring growth.

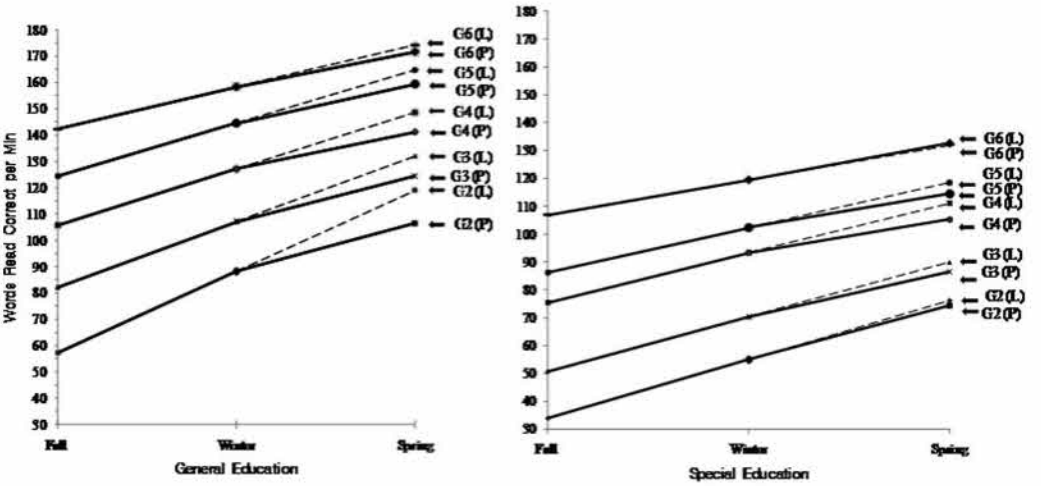


Figure 1. Differences for linear (L, solid line) and piecewise (P, dashed line) models of growth (G) for students in second through sixth grade within either the general (left) or special education population (right)

CBM-R growth was modeled and analyzed for five primary grades (second through sixth) and two student populations within and across seasons.

The results of this study support four general conclusions. Descriptive data make clear that, first, annual growth is greater among students in the general education population than for those in the special education population, especially in the fall season; and, second, more growth occurs in the early grades with less in the upper grades. Modeling and inferential analysis provide support for the third and fourth conclusions that follow. That is (third conclusion), there is a significant *seasonal effect* such that more growth occurs in the fall season than in the spring season for general education, but the magnitude of that effect declines with each progressive grade level. Finally, the seasonal effect is less pronounced among students in the special education population. In general, linear growth should not be assumed, seasonal effects should be expected—at least within the general education population—and piecewise models tend to fit better than linear models of growth.

Grade and Population Effects

Larger magnitudes of growth were observed for students in general education as compared to special education. This outcome was observed across five districts where CBM-R was used to screen, monitor, and evaluate instructional effects as part of a problem solving model. The systematic and ongoing use of CBM-R for these purposes did not establish equivalent growth rates between the student populations. Despite inconsistent growth rates across student groups, the rate of growth among students in the special education population was greater than observed in other studies (Deno et al., 2001; Fuchs et al., 1993). For example, the most recent estimates for average growth among students in the special education population were ~0.60 WRCM/week (Deno et al., 2001). That rate of growth was relatively consistent across grades. In contrast, the average growth estimates among students in the sample of students in special education used in this study were within the range of 0.67–1.17 WRCM/week (Table 3). Although cross-study comparisons cannot be used to establish that differences in systems-level practices (e.g., problem solving) and instructional practices were causally related to

the observed differences in growth rates, future research should investigate the phenomena more closely.

The results of this study converge with previous research (Deno et al., 2001; Fuchs et al., 1993) to support the conclusion that there is more CBM-R growth within the early primary grades. This phenomenon was consistent for both fall and spring seasons in both the general education and special education populations (Table 1). There was a steady decline for both general education and special education in weekly growth as grade level was increased. There was a corresponding decline in the magnitude of seasonal effects as grade levels increased. Results also support the conclusion that seasonal effects are more robust for students served within general education and less robust for students served within special education.

Seasonal Effect

The primary purpose of this research was to determine whether the rate of CBM-R growth is consistent across the year. Notwithstanding limitations, a seasonal effect was sufficiently robust and consistent across grades and student populations to warrant serious consideration. Estimates of weekly CBM-R growth within season are reported in Table 1. They are also depicted in Figure 1. Both visual analysis and statistical analysis of growth models support the conclusion that there was a *difference in the average rate of growth across seasons*, and that difference occurred in all grades, but was generally more robust in the lower grades of general education (Figure 1). With few exceptions, more growth occurred in the fall than in the spring, which is depicted as the differences in the slopes within Table 3.

The *magnitudes of seasonal effects were inconsistent* across grades and student populations. In general education, the difference between fall and spring slopes ranged from 0.14 to 0.69 per week, and all the differences were statistically significant with substantial effect sizes (Table 3). That effect is illustrated in Figure 1, which reveals that the interactions between seasonal effect and student population were most distinct in second grade. That is, the seasonal effect was most robust for the sample of students in second-grade general education classes, with a difference in growth from fall to winter of 0.69, and minor in comparison for students in second-grade special education classes, with a difference of 0.09.

The results of this study support the conclusion that growth differs between fall and spring seasons, which is inconsistent with a number of currently held views in the field. The consistency of the results across grades for general education begs the question of why this effect may be occur and what are the potential implications. There are a number of alternate hypotheses that offer some insight on the phenomenon and directions for future research.

Explanatory Hypotheses

This study was designed to examine *whether* seasonal effects are likely to occur in cases of CBM-R tri-annual assessment and challenge the assumption of linear growth. Indeed, the results do provide support for seasonal effects; however, this study was not designed to examine the cause of seasonal effects. That is left to future research to isolate the cause(s) and potential remedies to accelerate reading achievement within seasons with deficit growth. The result of both this study and Ardoin and Christ (2008) converge to provide evidence that the seasonal effect is likely to coincide with deficit levels of growth in the latter part of the academic year, or the spring season, especially for the students in the early primary general education population. It seems that changes in instructional conditions, or student motivation, might explain deficit rates of growth within the spring season unless there is an active intervention to prevent those changes. It seems that data-based decision making could function to reduce, or reverse, the pattern of seasonal effects that were replicated across this study and that of Ardoin and Christ.

Teacher expectations. There are a number of ways in which teacher expectations can have an influence on student behavior,

which might be conveyed through discrepant use of praise and punishment statements, content-relevant feedback, individualized attention, differentiated instruction, and variability in the opportunities to respond and engage in instructionally relevant practice (Good, 1981). Prior research yielded conclusions that there are indeed sometimes differences in teacher treatment of high- and low-achieving students over the course of the year, which affected the amount and type of instruction and feedback that students received at different points in the academic year (Good et al., 1980). Researchers concluded that teachers are more invested in influencing and shaping student behavior at the beginning of the school year than later in the year. Furthermore, the differentiation in teacher behavior may have sustained effects and become more influential as the school year progresses, resulting in lower achievement and growth later in the school year in low-achieving students. Other research also demonstrated the potential influence of teacher expectations, including expectation biases across ethnic groups (McKown & Weinstein, 2007) and ability groups (Rosenthal & Jacobson, 1966).

Classroom management. There is evidence that classroom management and organizational strategies are implemented with greater fidelity and effort in the early part of the academic year (Cameron, Connor, & Morrison, 2004). The same research provides evidence that that those educators who approach classroom management and organization with the greatest fervor early in the year are those same educators whose practices drop off most substantially later in the academic year. Improved levels of engagement and instructional gain are clearly associated with good classroom management and organization, and are among better predictors of student achievement (Brophy, 1987). Taken together, research provides a foundation to postulate that differential levels of classroom management and organization across seasons might influence student achievement. This effect is hypothesized to be the greatest at lower grade levels, as younger students require greater direction in organization practices (Cameron et al., 2004). The pattern of seasonal effects observed in this study coincides with a classroom management and organizational hypothesis.

Intervention to accelerate seasonal growth. The results of this study indicate that the seasonal effect with deficit growth in the spring season was most pronounced in the middle primary grades (second and third) and were minimized for the special education population. The results of Graney et al. (2009) indicate that the seasonal effect can function so that there is deficit growth in the fall season and accelerated growth in the spring. As noted in the Introduction, the data set that was analyzed by Graney et al. (2009) derived from a university and school partnership to implement a data-based decision making and evidence-based interventions. Those factors might have influenced the patterns of seasonal effects. There is a substantial literature base to support the conclusion that data-based decision making can function to accelerate academic growth (Carnine & Granzin, 2001; Fuchs & Fuchs, 1986; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Fuchs, Fuchs, Hamlett, & Whinnery, 1991; Graney et al., 2009), and there is substantial theoretical support for the use of evidence-based practices, which were both components of service delivery that were targeted by the partnership program. It might be that progress monitoring and data use accelerated learning and ameliorated the deficit growth associated with seasonal effects. Indeed, students who are served within special education have explicit goals and their progress is monitored. Although it is not possible to draw any final conclusions, it may well be that the lack of convergence in seasonal patterns between Graney et al. and other studies—along with minimal seasonal effects observed for the special education population—provides further evidence for the use of data-based decision making to improve growth during otherwise deficit periods.

Limitations and Future Research

The sample studied was entirely from five relatively ethnically homogeneous dis-

tricts in a single midwestern state. Future research should attempt to corroborate these findings across a more diverse sample of students and educational programs. Finally, the results of this study derive from data collected at three points in the academic year (fall, winter, spring). Future research should examine whether the same conclusions would result from the analysis of a more robust data set (e.g., weekly, monthly) and with more precise dates that define the data collection schedule. Graney et al. (2009) correctly criticize the methods of this study and that of Ardoin and Christ (2008) because the precise dates of data collection were not used in the analysis; rather, both studies rely on data that are centered by month and not day. Although the authors of this and the prior study believe the conclusions would be substantially similar with different centering methods or more precise dates, the criticism ought to be accounted for in future research. Future research should investigate the nature of growth across summative and formative measurement conditions. Research should also examine the effect on educational decisions when the potential for seasonal effects is ignored and consistent annual growth is assumed.

Implications for School-Based Decision Making

CBM-R was developed to guide routine instructional decisions and enhance instructional effects (Deno, 1986), and it has emerged as one of the primary methods of assessment to guide problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention models of service delivery (National Association of State Directors of Education, 2005). As described in the Introduction of this study, CBM-R is used to establish expectations for both the level and rate of reading achievement throughout the primary grades; therefore, the results of this study provide at least two clear implications. First, there is good reason to question the validity of linear growth estimates as criteria to guide progress monitoring (i.e., discrepant in rate) and dual-discrepancy decisions (i.e., discrepant in both level and rate; Fuchs, Fuchs, McMaster & Al Otaiba, 2003). Second, seasonal effects can be either admired as a phenomenon or targeted for intervention. These implications are addressed in what follows.

First, if CBM-R estimates of growth are used to evaluate instructional effects and response to intervention, then those who interpret data should be aware that the magnitude of growth might be less in the spring than in the fall. There are two weighty implications that derive from this study. First, a dual-discrepancy approach to special education eligibility, which compares the observed level and rate of achievement to criterion standards (Fuchs, Fuchs, McMaster & Al Otaiba, 2003), might function to overidentify students during deficit seasons of growth and underidentify students during accelerated seasons of growth. If the data follow the pattern observed in this study, then it is likely that the dual-discrepancy model might underidentify students in the fall season and overidentify students in the spring season. That consequence is counter to the values of early intervention and prevention inherent within problem solving and response to intervention. Researchers and practitioners should pay close attention to the patterns of growth and the potential for seasonal effects to influence decisions at the local level. Accelerated rates of growth should be expected in the fall.

Second, the pattern of seasonal effects that are observed across populations and studies provide preliminary evidence that goal setting, progress monitoring and data-based decision making can minimize, eliminate or reverse the pattern of seasonal effects (e.g., Graney et al., 2009). Although future research is necessary, it is a working thesis that the systematic use of data to evaluate instructional effects and select effective instructional strategies effectively accelerates learning and prevents problems at the individual, group, and systems level. After all, data-based decision making is a central tenet of problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention (National Association of State Directors of Education, 2005); however, data must be both systematically collected *and* sys-

tematically used to establish effects (Fuchs et al., 1992, 1991; Fuchs & Fuchs, 1986). The mere act of data collection without data use typically fails to establish effects (Fuchs & Fuchs, 1986). The explanatory hypotheses proposed in the previous section establish that deficit growth in the spring season might be remediated at the systems level if data are systematically collected and systematically used. If detected, this problem (seasonal effect) is ripe for problem analysis at the local level. Deficit growth within a single season should not be admired as an acceptable phenomenon, but, instead, it should be targeted for analysis and intervention.

Christ (2008) defined problem analysis as the "collection, summary, and use of information to systematically test, reject, or verify relevant hypotheses to establish problem solutions" (p. 159). That is, deficit rates of growth in the spring season might be remediated if expectations remain high and data are used to evaluate and maintain high standards for instructional effects. There is no reason to assume that the seasonal effect is inherent to the educational system. Instead, hypotheses related to the causal and maintaining variables should be developed—such as those presented earlier—and tested through intervention to remediate the problem. It is likely that the pattern of seasonal effects is causally related to what occurs in the classroom; and what occurs in the classroom can be influenced by the systematic used of data to evaluate of response to intervention for universal, supplemental, and intensive tiers of service delivery. This study provides evidence of seasonal effects as a potential target for intervention at the system, grade, or classroom level.

References

AIMSweb. (2002). *Reading curriculum-based measurement (R-CBM) standard oral reading fluency assessment passages.* Retrieved October 2003, from http://www.aimsweb.com

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum based measurement slope estimate using data from tri-annual universal screenings *School Psychology Review, 37,* 109–125.

Bollman, K. A., Silberglitt, B., & Gibbons, K. A. (2007). The SCRED Model: Incorporating systems-level organization and a multi-tiered problem-solving process for intervention delivery. In M. Burns, S. Jimerson, & A. VanDerHayden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 321–332). New York: Springer.

Brophy, J. (1987). Synthesis of research on strategies for motivating students to learn. *Educational Leadership, 45*(2), 40–48.

Cameron, C. E., Connor, C. M., & Morrison, F. J. (2004). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology, 43,* 61–85.

Carnine, D., & Granzin, A. (2001). Setting learning expectations for students with disabilities. *School Psychology Review, 30,* 466–472.

Christ, T. J. (2008). Best practices in problem analysis. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 159–176). Bethesda, MD: National Association of School Psychologists.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74,* 68.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219–232.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15,* 358–374.

Deno, S. L. (2002). Problem solving as best practices. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 37–56). Bethesda, MD: National Association of School Psychologists.

Deno, S. L. (2005). Problem solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 10–42). New York: Guilford Press.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30,* 507–524.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). An introduction to latent variable growth curve modeling: Concepts, issues, and applications (2nd ed.). Mahwah, NY: Lawrence Erlbaum Associates.

Flora, D. B. (2008). Specifying piecewise latent trajectory models for longitudinal data. *Structural Equation Modeling, 15,* 513–533.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53,* 3–14.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10,* 43–52.

Fuchs, L. S., Fuchs, D., & Stecker, P. M. (1989). Effects of curriculum-based measurement on teachers' instructional planning. *Journal of Learning Disabilities, 22,* 51–59.

Fuchs, L. S., & Shinn, M. R. (1989). Writing CBM IEP objectives. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 131–152). New York: Guildford Press.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58,* 436–450.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22,* 27–48.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Whinnery, K. (1991). Effects of goal line feedback on level, slope, and stability of performance within curriculum-based measurement. *Learning Disabilities Research and Practice, 6,* 66–74.

Fuchs, D., Fuchs, L. S., McMaster, K. N., & Al Otaiba, S. (2003). Identifying children at risk for reading failure: Curriculum-based measurement and the dual-discrepancy approach. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 431–449). New York: Guilford Press.

Good, T. L. (1981). Teacher expectations and student perceptions: A decade of research. *Educational Leadership, 38*(5), 415–422.

Good, T., Cooper, H., & Blakey, X. (1980). Classroom interaction as a function of teacher expectations, student sex and time of year. *Journal of Educational Psychology, 72*, 378–385.

Graney, S. B., Missall, K. N., Martinez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School* Psychology, 47, 121–142.

Howe, K. B., & Shinn, M. M. (2002). Standard *reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Retrieved October 2004, from http://www.aimsweb.com

Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education for Students Placed at Risk, 9,* 169–188.

McKown, C., Weinstein, R. S. (2007). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*, 235–261.

National Association of State Directors of Education. (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: Author.

National Reading Panel. (2000). *Report of the National Reading Panel Subgroups: Teaching children to read* (No. NIH 00–4754). Bethesda, MD: National Institute of Child Health and Human Development.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports, 19,* 115–118.

Satorra, A., & Saris, W. E. (1985). The power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50*, 83–90.

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education, 34,* 164–172.

Shinn, M. R. (1988). Development of curriculum-based local norms for use in special education decision-making. *School Psychology Review, 17,* 61–80.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children.* New York: Guildford Press.

Shinn, M. (2008). Best practices for using curriculum based measurement in a problem solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 243–262). Bethesda, MD: National Association of School Psychologists.

Stage, S. A. (2001). Program evaluation using hierarchical linear modeling with curriculum-based measurement reading probes. *School Psychology Quarterly, 16,* 91–112.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41,* 85–120.

Date Received: May 30, 2007
Date Accepted: April 21, 2010
Action Editor: Tanya Eckert ■

Theodore J. Christ, PhD, is an associate professor in the Educational Psychology Department at the University of Minnesota in Minneapolis—Saint Paul. He is the codirector of the Research Institute for Problem Solving. He is engaged in research and training to evaluate, develop, and enhance the tools and procedures that enable data-based decision making and problem solving within school settings. Interests include progress monitoring, curriculum-based measurement, direct behavior ratings, and computer adaptive testing.

Benjamin Silberglitt completed his PhD in Educational Psychology at the University of Minnesota at Minneapolis—Saint Paul. His areas of interest include student assessment school organization and system-level change, problem solving, response to intervention, and data-based decision making. He is director of Software Applications for TIES, a technology cooperative of 39 Minnesota school districts.

Seungsoo Yeo, PhD, is a researcher at Korean Educational Development Institute. He received his PhD from the University of Minnesota at Twin Cities in 2008. His primary research interests include curriculum-based measurement, response to intervention, and learning disabilities. His methodological work has focused on latent growth modeling and multilevel meta-analysis.

Damien C. Cormier, MA, is a school psychology doctoral student in the Educational Psychology Department at the University of Minnesota. He is a graduate research assistant at the National Center on Educational Outcomes. His research interests are in the areas of large-scale assessment policy and practice, measurement issues in individual assessment, high-functioning autism, and cultural and linguistic biases in assessment.

Copyright of School Psychology Review is the property of National Association of School Psychologists and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

I was mistaken when I indicated that the Reading Fluency Monitor was developed by Read Naturally. It was actually developed by the RMC Research Corporation out of Portsmouth, New Hampshire.

Sorry for the misunderstanding,

- Dave

David Heistad, Ph.D., Executive Director
Research and Evaluation Department
Bloomington Public Schools - ISD 271
1350 W. 106th Street
Bloomington, MN 55431
ph: 952-681-6485
cell: (b)(6)

Karla Ramy

From:Karla Ramy
Sent:26 Aug 2013 15:35:58 +0000
To:What Works
Subject:Read Naturally - Tom Ihnot
Attachments:Christ Silberglitt Yeo Cormier (2010).pdf, Heistad.docx

Jill Constantine,

I am requesting that What Works Clearinghouse (WWC) change its Beginning Reading effectiveness rating on reading fluency for Read Naturally, Inc. from mixed effects to positive effects for several reasons. Please consider and respond to the following information.

First, in a recent WWC review of a study by Dr. David Heistad, statistically significant gains in oral reading fluency for third grade students across multiple schools were not considered in the evaluation of the effects of a Read Naturally intervention on fluency. After reviewing the most recent WWC Procedures and Standards Handbook (February, 2013), the reason for exclusion may be due to a simple misunderstanding.

Page 15 of the WWC Procedures and Standards Handbook (Outcome Eligibility and Reporting section) states the following:

"….outcome measures is that they not be over aligned with the intervention. When outcome measures are closely aligned with or tailored to the intervention, the study findings may not be an accurate indication of the effect of the intervention. For example, an outcome measure based on an assessment that relied on materials used in the intervention condition but not in the comparison condition…."

The measure of oral reading fluency in Dr. Heistad's study was *The Reading Fluency Monitor*. It should be noted that *The Reading Fluency Monitor* is not a component of the Read Naturally Intervention. The *Reading Fluency Monitor* was developed by RMC Research Corporation, located in Portsmouth, New Hampshire. Dr. Heistad's implication that *The Reading Fluency Monitor* was developed by Read Naturally, Inc. was a miscommunication in the Test Instruments section of his study. Heistad's verification of this misunderstanding is attached.

Second, the Arvans study summary analysis underestimated the effect of the Read Naturally group on oral reading fluency. The Read Naturally group had a large effect size of .81 for fluency. The control EAU group had a moderate effect size of .57 for fluency. Both groups had excellent fluency gains. It should be noted that the use of the Read Naturally Intervention was extensive at the school before the study was conducted. The teachers were well versed in the Read Naturally strategy. Students used Read Naturally in the first and second grade in the two years prior to the study. Also, some students in the EAU group used Read Naturally during the study. That is, it seems as though the EAU group may not have been demonstrative of a true control group. The difference between the Read Naturally group's .81 effect size and EAU's .57 effect size was .24 after eight weeks. An effect size difference of .24 is significant when extrapolated over a school year.

Ethan R. Van Norman, M.A., performed a similar analysis to what appeared in the Christ and Davie study (2008). In that study, the authors first calculated a slope estimate from three time points for each student in the EAU and the experimental group. The slope estimate represented the number of words read correct per minute (WCPM) improvement per week. The mean and standard deviation of slopes

were then calculated for each group. The percent of improvement of the experimental group in relation to the EAU group was calculated. After this, the authors used the percent of improvement and applied it to an aggressive rate of growth (1.50 WCPM improvement per week). That value was then multiplied by 36, and 1.50 was multiplied by 36. The difference between those values was then interpreted as a hypothetical effect if the Read Naturally Intervention was delivered across an entire school year.

A similar analysis was conducted on the Arvans dataset. Slope estimates were calculated for each student from two observations eight weeks apart. The mean slope value for the experimental group was 2.92 WCPM improvement per week (SD = 1.54) compared to the EAU group, which had a mean slope estimate of 2.24 (SD = 2.36). The .68 difference in mean slope for the Read Naturally Intervention group represents a 30% improvement over the EAU group. Assuming an aggressive rate of growth of 1.50 WCPM for typical students, a 30% increase would translate to a 1.95 rate of growth. Extended across 36 weeks, this represents a net increase of 70 WCPM compared to 54 WCPM with a typical rate of growth (1.50 WCPM improvement per week across 36 weeks).

Although not ideal, slope estimates from two time points have been used to summarize growth in previous CBM-R research studies (e.g., Christ, Silberglit, Yeo & Cormier, 2010). See the attached document for further details.

Another way to analyze the Arvans study is to use the Hasbrouck-Tindal Oral Reading Fluency Norms (http://www.readnaturally.com/howto/orftable.htm). The table suggests that third grade students at the 50th percentile have an average weekly improvement of 1.1 WCPM. The performance of the Read Naturally group in the Arvans study is also impressive when compared to the typical rate of fluency growth of third grade students.

Ethan R. Van Norman, M.A., did an analysis to extrapolate growth in the same manner using normative values. For third grade students, weekly growth estimates for students in the 50th percentile typically approximate 1.10 WCPM per week. Assuming the Read Naturally group's 30% improvement over the EAU group and the intervention is delivered for 36 weeks, a student in the 50th percentile, on average would be improving at a rate of 1.43 WCPM per week. After 36 weeks this would translate to an improvement of 51 WCPM compared to 40 WCPM if the student did not receive the intervention. This is a substantial difference.

Considering all of this information, Read Naturally programs should receive a positive effects rating for fluency. A timely reevaluation is critical so that WWC may change Read Naturally's effectiveness rating for fluency from mixed effects to positive effects. Please respond to the points I've raised in this email and let me know if you need further information.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally®
CELEBRATING
20 YEARS

Julia Lyskawa

From:Julia Lyskawa
Sent:27 Aug 2013 17:08:45 -0400
To:WhatWorks
Subject:FW: What Works Clearinghouse (QR2010013)
Attachments:QRT 2010013 Response_8.27.2013.pdf, QRT2010013_Response 5.13.2013.pdf, QRT2010013_Response_2.26.2013.pdf

---

From: What Works
Sent: Tuesday, August 27, 2013 5:08 PM
To: 'kramy@readnaturally.com'; 'info@readnaturally.com'
Subject: What Works Clearinghouse (QR2010013)

Dear Mr. Ihnot:

Attached in the document entitled *QRT 20100013_Response_8.27.2013* is a response to the questions you raised in your inquiries concerning the What Works Clearinghouse reviews of studies of Read Naturally®. The letter makes reference to two other letters to you from the What Works Clearinghouse, and those are attached as well.

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/.

August 26, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013, WWC2304

Dear Mr. Ihnot,

Thank you for your continued correspondence regarding the What Works Clearinghouse (WWC) reports that discuss Read Naturally, Inc. products and the Read Naturally® strategy. This letter is in response to your correspondence of June 24, July 17, August 5, and August 7 of this year.

Program descriptions. You previously raised concerns that it was difficult to distinguish between Read Naturally, Inc. and the Read Naturally® strategy and products. The WWC indicated in the May 12 response that revisions to the product description section would be made. The Beginning Reading report was updated in July, and changes to the product description in the other three intervention reports are under way. Similarly, we are adding additional detail, if provided in the original publication, to clarify which Read Naturally, Inc. product was used and whether key elements were excluded in the implementation.

Order of findings. We appreciate and agree with your suggestion of July 17, in which you propose that the WWC list findings in order of their strength, consistent with the presentation on the web. Therefore, we will revise the order of findings listed in the text on the front page of the report, the associated table, and the website summary for the three reports with findings in more than one domain. We anticipate that the revised intervention reports will be posted to the web by the end of September.

No discernible effect. Two of your concerns relate to the use of the phrase "no discernible effect." Based on analyses (of eligible outcomes) that meet standards, the WWC uses a set of rules to characterize study findings within a domain (see pages 96 and 97 of the Procedures and Standards Handbook, version 2.1). Next, the WWC combines these study-level findings to determine an intervention rating (see pages 23 and 24 of the Handbook). If "none of the studies shows statistically significant or substantively important effects, either positive or negative," the WWC assigns a rating of effectiveness of "no discernible effects" for that domain.

Study-specific issues. With respect to your other concerns regarding study inclusion and implementation, we have previously addressed these issues in our responses of February 26 and May 12, attached.

I hope that this letter and the forthcoming changes to reports address your concerns. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov.

Sincerely,

(b)(6)

Jill Constantine
Director
What Works Clearinghouse

May 12, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013, WWC2304, and WWCPC3661

Mr. Ihnot:

Thank you for your response to Dr. Constantine's letter of February 26, 2013. The What Works Clearinghouse (WWC) Quality Review Team has reviewed your correspondence dated March 5, March 6, and March 20, 2013.

Your correspondence raises five questions with respect to reviews of Read Naturally®: (1) results when the term "Read Naturally" is searched using Google; (2) the inclusion of five studies (Hancock 2002; Kemp 2006; Denton et al. 2006; Chenault 2004; and Chenault et al. 2006); (3) the update to the Beginning Reading intervention report on Read Naturally®; (4) reviews by the National Center on Intensive Intervention (NCII) of studies of Read Naturally®; and (5) the difficulties an individual might have in distinguishing between Read Naturally, Inc. and Read Naturally®. Below we provide responses to these questions, including a brief summary of our reasoning in each area. The documents explaining all of our procedures and review criteria continue to be publically available on the WWC website. As suggested in our previous letters, please refer to the *WWC Procedures and Standards Handbook* and the topic area review protocol for a particular intervention report.

Google Results for "Read Naturally"

WWC does not have any control over or input as to what results are presented by Google or any other search engine when an individual searches "Read Naturally." It is our understanding that the various search engines use algorithms that generally are proprietary.

Inclusion of Hancock (2002), Kemp (2006), Denton et al. (2006), Chenault (2004), and Chenault et al. (2006)

We appreciate the information you provided to us in Addendums 1 and 2 of your March correspondence. Our response is the same as in our February 26 letter: WWC followed protocol in determining whether studies supported an examination of the effectiveness of Read Naturally®. To assist the reader in understanding what was being examined, WWC includes in the intervention report a description of both the intervention and comparison conditions to the extent they are described in each individual study. This includes describing whether elements of an intervention were not implemented. WWC reviews publicly available research, and these reports may differ in the amount of detail provided on the condition of interest to WWC. To the extent possible, the reports summarize what is known about how a particular intervention was implemented.

Update to the Beginning Reading Intervention Report on Read Naturally®

WWC is currently updating the Beginning Reading intervention report for Read Naturally®. As the developer and distributor of Read Naturally®, a courtesy copy of the embargoed intervention report will be shared with you 24 hours before release. We cannot provide details on the content of the update before you receive the copy, which is embargoed until it is released. All of the studies you listed in your March 20 email were reviewed as part of the updated report.

NCII Reviews of Read Naturally® Studies

Eight of the nine studies that you identified in Addendums 1 and 2 were reviewed for the update to the Beginning Reading intervention report on Read Naturally® (Hesitad, Christ, Tucker, and five case studies). As the report has not yet been released, we cannot discuss individual study ratings or the overall summary of effectiveness. The remaining case study (Case 5, "Throwing Sixth Graders a Lifeline") is not within the age range of the Beginning Reading review and thus is not eligible for review in a Beginning Reading report, although it could be eligible for review in the Adolescent Literacy report on Read Naturally®.

Distinguishing Between Read Naturally, Inc. and Read Naturally®

The program details, which are included in the main part of the intervention report, are shared with developers or distributors for review. We provided you with that description, and you provided input in February 2012. However, we agree that it is important and useful to readers to distinguish among the different products that support Read Naturally® strategies. Thus, we will revise the product description further to discuss Read Naturally® as a strategy supported by the four products (Read Naturally® Live, Read Naturally® Encore, Read Naturally® ME, and Read Naturally® SE) that can be purchased from Read Naturally®, Inc. To the extent that the specific products are described in each study, we will also clearly note which products were used. In addition, we will use the copyright symbol (®) with the phrase Read Naturally® throughout the study.

WWC appreciates your careful review of our products and processes. If you or another member of Read Naturally, Inc. are interested in learning more about WWC's review standards, please consider applying for one of the training sessions offered by the Institute of Education Sciences. You can send an email to the WWC help desk asking to be added to the email list for training announcements.

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

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

February 26, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013

Dear Mr. Ihnot:

Thank you for your inquiries concerning the What Works Clearinghouse (WWC) reviews of four studies of Read Naturally®—Kemp (2006), Denton et al. (2004), Hancock (2002), and Chenault et al. (2006)—and for your email regarding my response to Dr. Kevin Feldman's inquiry about the WWC reviews of studies of Read Naturally®. In response to your concerns, we conducted an independent quality review. The WWC quality review team responds to concerns raised by study authors, curriculum developers, or other relevant parties about WWC reviews published on our website. When a quality review is conducted, a researcher who was not involved in the initial reviews undertakes an independent assessment of the studies in question. The researcher also investigates the procedures used and decisions made during the original reviews of the studies. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the WWC.

The quality review considered the concerns from your emails regarding study inclusion, implementation, student sample, and outcomes.

Study Inclusion

The WWC followed protocol in choosing to review these four studies. The WWC screens studies based on a number of factors, including relevancy and methodology criteria (as described in the WWC Procedures and Standards Handbook, pages 8–10, available at http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19). The WWC does not screen based on whether the author(s) explicitly intended the study as an evaluation of an intervention, but rather whether the study presents a primary analysis of the effect of an intervention. This screening process allows for the review of a wide range of relevant and methodologically sound studies.

Thank you for bringing to my attention the studies by Mesa (2004) and Tucker (2010). The Tucker (2010) study will be included in our literature search for updating the Intervention Report

for Read Naturally®, reviewed under the Beginning Reading Evidence Review Protocol. In addition, as you note, the Mesa (2004) study was included in the original review of Read Naturally®, but will be reviewed again in an updated review.

Implementation

The WWC followed protocol in the manner in which the four studies are described in WWC publications. The WWC Handbook says, "The WWC makes no adjustments or corrections for variations in implementation of the intervention; however, if a study meets standards and is included in an intervention report, descriptions of implementation are provided in the report appendices to provide context for the findings" (page 16). This approach is appropriate because there is no standard metric for fidelity to intervention design. Thus, the WWC includes studies with variation in fidelity and does not evaluate implementation fidelity. The quality review team verified that variations in implementation that are noted in the four studies and that may affect the interpretation of findings were properly included in the WWC publications.

Regarding the Kemp (2006) study, your emails raised a concern about implementation of the Read Naturally® program. The quality review noted that the study does not provide any indication of deviation from program design. Specifically, the study states, "… it could be concluded that all teachers implemented the Read Naturally® program as prescribed" (page 40). Based on this information, the quality review concluded that there was no evidence that deviations in implementation should have been noted in the WWC Intervention Report. Regarding your specific concern about the teacher-to-student ratio, the study does not note any deviation from the Read Naturally® program in this area and does not provide any information on the ratio.

You also expressed concern over deviations in the implementation of Read Naturally® in the Denton et al. (2004) study. The quality review found that the inclusion of additional activities was noted in the study and reported in the WWC Intervention Report (see footnote 7 on page 3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). The other potential deviations that you noted (accountability to pass criteria, use of audio tapes, and the number of sessions per week) were not identified in the study and are therefore not noted in the WWC publication.

Regarding your concerns that the Read Naturally® strategy was not implemented fully for the Hancock (2002) study, the quality review found that WWC procedures were followed in the WWC Intervention Report. The WWC contacted Hancock regarding the implementation of Read Naturally® in the 2002 study. Her response indicated that the study excluded Read Naturally®'s pre-reading vocabulary instruction component and the placement system to individualize instruction. The WWC Intervention Report was revised to note these deviations in implementation (see footnote 4 on page 2, available at http://ies.ed.gov/ncee/wwc/InterventionReport.aspx?sid=407). Other potential concerns with implementation described in your email were not reported by the author and were not included in the revised WWC Intervention Report.

The quality review also investigated concerns that you raised in prior emails. The email from June 2012 regarding Dr. Feldman's questions was addressed in my letter to Dr. Feldman; I will not

repeat the quality review findings here. Your email from May 2012 included some text that was reported to be a quote from Hancock's email to the WWC, but the quoted statements were not part of Hancock's email text. The misreported quote is also found on the Read Naturally® website at http://www.readnaturally.com/approach/hancock.htm.

Regarding the Chenault et al. (2006) study, you expressed concern with the implementation of Read Naturally® related to the selection of students and the limited number of sessions. The quality review found that the WWC Intervention Report appropriately notes both of these factors (see pages 2–3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=409).

Student Sample

You also raised a concern that the Kemp (2006) study included a large share of students who were above the 50th percentile in reading fluency who were not "at-risk" readers. The quality review verified that the WWC Intervention Report accurately states that the study population for the results reported by the WWC was third-grade English language learner (ELL) students (page 2, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). Furthermore, the quality review noted that Read Naturally® program materials state that the program "… products help students of all ages and abilities—whether special education, ELL, Title I, or mainstream—to improve their reading skills, gain confidence, and transform their attitudes toward school and learning" (available at http://www.readnaturally.com/products/default.htm).

Finally, regarding the differences in pre-test performance among native-speaking students in the treatment and control groups, the quality review found that these concerns are not relevant to this WWC publication, which reports on findings only for English language learner students. Furthermore, the quality review concluded that the WWC appropriately restricted the reported findings to impacts assessed relative to the comparison group, and thus did not include achievement gains in oral reading fluency.

Outcomes

Regarding the Denton et al. (2004) study, you expressed concern that the study did not assess fluency. The quality review noted that the WWC Intervention Report clearly states that the study assessed reading achievement (page 4, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408), an outcome of relevance to the WWC ELL topic area. The quality review noted that the Read Naturally® program materials state there are several components of the program beyond fluency that are related to reading achievement: "Our industry-leading programs develop and support the five essential components of reading, identified by the National Reading Panel: phonemic awareness, phonics, fluency, vocabulary, and comprehension" (see "The Read Naturally Strategy," available at http://www.readnaturally.com/approach/default.htm).

I hope that this letter has addressed your concerns. You requested that we meet to discuss the WWC reviews of Read Naturally®. However, WWC policy requires all correspondence in writing. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov.

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

cc: Joy Lesnick, IES

WhatWorks

From:WhatWorks
Sent:28 Aug 2013 13:40:22 +0000
To:'kramy@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot
Dear Mr. Ihnot,

Thank you for your email. The QRT will send a separate response to your email of August 26, 2013.

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Monday, August 26, 2013 11:36 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Jill Constantine,

I am requesting that What Works Clearinghouse (WWC) change its Beginning Reading effectiveness rating on reading fluency for Read Naturally, Inc. from mixed effects to positive effects for several reasons. Please consider and respond to the following information.

First, in a recent WWC review of a study by Dr. David Heistad, statistically significant gains in oral reading fluency for third grade students across multiple schools were not considered in the evaluation of the effects of a Read Naturally intervention on fluency. After reviewing the most recent WWC Procedures and Standards Handbook (February, 2013), the reason for exclusion may be due to a simple misunderstanding.

Page 15 of the WWC Procedures and Standards Handbook (Outcome Eligibility and Reporting section) states the following:

"….outcome measures is that they not be over aligned with the intervention. When outcome measures are closely aligned with or tailored to the intervention, the study findings may not be an accurate indication of the effect of the intervention. For example, an outcome measure based on an assessment that relied on materials used in the intervention condition but not in the comparison condition…."

The measure of oral reading fluency in Dr. Heistad's study was *The Reading Fluency Monitor*. It should be noted that *The Reading Fluency Monitor* is not a component of the Read Naturally Intervention. The *Reading Fluency Monitor* was developed by RMC Research Corporation, located in Portsmouth, New Hampshire. Dr. Heistad's implication that *The Reading Fluency Monitor* was developed by Read Naturally, Inc. was a miscommunication in the Test Instruments section of his study. Heistad's verification of this misunderstanding is attached.

Second, the Arvans study summary analysis underestimated the effect of the Read Naturally group on oral reading fluency. The Read Naturally group had a large effect size of .81 for fluency. The control EAU group had a moderate effect size of .57 for fluency. Both groups had excellent fluency gains. It should be noted that the use of the Read Naturally Intervention was extensive at the school before the study was conducted. The teachers were well versed in the Read Naturally strategy. Students used Read

Naturally in the first and second grade in the two years prior to the study. Also, some students in the EAU group used Read Naturally during the study. That is, it seems as though the EAU group may not have been demonstrative of a true control group. The difference between the Read Naturally group's .81 effect size and EAU's .57 effect size was .24 after eight weeks. An effect size difference of .24 is significant when extrapolated over a school year.

Ethan R. Van Norman, M.A., performed a similar analysis to what appeared in the Christ and Davie study (2008). In that study, the authors first calculated a slope estimate from three time points for each student in the EAU and the experimental group. The slope estimate represented the number of words read correct per minute (WCPM) improvement per week. The mean and standard deviation of slopes were then calculated for each group. The percent of improvement of the experimental group in relation to the EAU group was calculated. After this, the authors used the percent of improvement and applied it to an aggressive rate of growth (1.50 WCPM improvement per week). That value was then multiplied by 36, and 1.50 was multiplied by 36. The difference between those values was then interpreted as a hypothetical effect if the Read Naturally Intervention was delivered across an entire school year.

A similar analysis was conducted on the Arvans dataset. Slope estimates were calculated for each student from two observations eight weeks apart. The mean slope value for the experimental group was 2.92 WCPM improvement per week (SD = 1.54) compared to the EAU group, which had a mean slope estimate of 2.24 (SD = 2.36). The .68 difference in mean slope for the Read Naturally Intervention group represents a 30% improvement over the EAU group. Assuming an aggressive rate of growth of 1.50 WCPM for typical students, a 30% increase would translate to a 1.95 rate of growth. Extended across 36 weeks, this represents a net increase of 70 WCPM compared to 54 WCPM with a typical rate of growth (1.50 WCPM improvement per week across 36 weeks).

Although not ideal, slope estimates from two time points have been used to summarize growth in previous CBM-R research studies (e.g., Christ, Silberglit, Yeo & Cormier, 2010). See the attached document for further details.

Another way to analyze the Arvans study is to use the Hasbrouck-Tindal Oral Reading Fluency Norms (http://www.readnaturally.com/howto/orftable.htm). The table suggests that third grade students at the 50th percentile have an average weekly improvement of 1.1 WCPM. The performance of the Read Naturally group in the Arvans study is also impressive when compared to the typical rate of fluency growth of third grade students.

Ethan R. Van Norman, M.A., did an analysis to extrapolate growth in the same manner using normative values. For third grade students, weekly growth estimates for students in the 50th percentile typically approximate 1.10 WCPM per week. Assuming the Read Naturally group's 30% improvement over the EAU group and the intervention is delivered for 36 weeks, a student in the 50th percentile, on average would be improving at a rate of 1.43 WCPM per week. After 36 weeks this would translate to an improvement of 51 WCPM compared to 40 WCPM if the student did not receive the intervention. This is a substantial difference.

Considering all of this information, Read Naturally programs should receive a positive effects rating for fluency. A timely reevaluation is critical so that WWC may change Read Naturally's effectiveness rating for fluency from mixed effects to positive effects. Please respond to the points I've raised in this email and let me know if you need further information.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

May 12, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013, WWC2304, and WWCPC3661

Mr. Ihnot:

Thank you for your response to Dr. Constantine's letter of February 26, 2013. The What Works Clearinghouse (WWC) Quality Review Team has reviewed your correspondence dated March 5, March 6, and March 20, 2013.

Your correspondence raises five questions with respect to reviews of Read Naturally®: (1) results when the term "Read Naturally" is searched using Google; (2) the inclusion of five studies (Hancock 2002; Kemp 2006; Denton et al. 2006; Chenault 2004; and Chenault et al. 2006); (3) the update to the Beginning Reading intervention report on Read Naturally®; (4) reviews by the National Center on Intensive Intervention (NCII) of studies of Read Naturally®; and (5) the difficulties an individual might have in distinguishing between Read Naturally, Inc. and Read Naturally®. Below we provide responses to these questions, including a brief summary of our reasoning in each area. The documents explaining all of our procedures and review criteria continue to be publically available on the WWC website. As suggested in our previous letters, please refer to the *WWC Procedures and Standards Handbook* and the topic area review protocol for a particular intervention report.

Google Results for "Read Naturally"

WWC does not have any control over or input as to what results are presented by Google or any other search engine when an individual searches "Read Naturally." It is our understanding that the various search engines use algorithms that generally are proprietary.

Inclusion of Hancock (2002), Kemp (2006), Denton et al. (2006), Chenault (2004), and Chenault et al. (2006)

We appreciate the information you provided to us in Addendums 1 and 2 of your March correspondence. Our response is the same as in our February 26 letter: WWC followed protocol in determining whether studies supported an examination of the effectiveness of Read Naturally®. To assist the reader in understanding what was being examined, WWC includes in the intervention report a description of both the intervention and comparison conditions to the extent they are described in each individual study. This includes describing whether elements of an intervention were not implemented. WWC reviews publicly available research, and these reports may differ in the amount of detail provided on the condition of interest to WWC. To the extent possible, the reports summarize what is known about how a particular intervention was implemented.

Update to the Beginning Reading Intervention Report on Read Naturally®

WWC is currently updating the Beginning Reading intervention report for Read Naturally®. As the developer and distributor of Read Naturally®, a courtesy copy of the embargoed intervention report will be shared with you 24 hours before release. We cannot provide details on the content of the update before you receive the copy, which is embargoed until it is released. All of the studies you listed in your March 20 email were reviewed as part of the updated report..

NCII Reviews of Read Naturally® Studies

Eight of the nine studies that you identified in Addendums 1 and 2 were reviewed for the update to the Beginning Reading intervention report on Read Naturally® (Hesitad, Christ, Tucker, and five case studies). As the report has not yet been released, we cannot discuss individual study ratings or the overall summary of effectiveness. The remaining case study (Case 5, "Throwing Sixth Graders a Lifeline") is not within the age range of the Beginning Reading review and thus is not eligible for review in a Beginning Reading report, although it could be eligible for review in the Adolescent Literacy report on Read Naturally®.

Distinguishing Between Read Naturally, Inc. and Read Naturally®

The program details, which are included in the main part of the intervention report, are shared with developers or distributors for review. We provided you with that description, and you provided input in February 2012. However, we agree that it is important and useful to readers to distinguish among the different products that support Read Naturally® strategies. Thus, we will revise the product description further to discuss Read Naturally® as a strategy supported by the four products (Read Naturally® Live, Read Naturally® Encore, Read Naturally® ME, and Read Naturally® SE) that can be purchased from Read Naturally®, Inc. To the extent that the specific products are described in each study, we will also clearly note which products were used. In addition, we will use the copyright symbol (®) with the phrase Read Naturally® throughout the study.

WWC appreciates your careful review of our products and processes. If you or another member of Read Naturally, Inc. are interested in learning more about WWC's review standards, please consider applying for one of the training sessions offered by the Institute of Education Sciences. You can send an email to the WWC help desk asking to be added to the email list for training announcements.

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

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

From: info@readnaturally.com
Sent: Wednesday, February 01, 2012 11:57 AM
To: What Works
Subject: Read Naturally - Tom Ihnot
Attachments: Intervention Summary for Read Naturally.doc; SCAN1267_000.tif

Dear Susanne James Burdumy,

I appreciate that What Works Clearinghouse (WWC) is updating its Beginning Reading Reviews. It provides an opportunity to correct a serious mistake. The current postings on Read Naturally cause confusion for educators and are a serious injustice to Read Naturally. I would appreciate an opportunity to explain why the Hancock and Denton studies should be removed from WWC website.

The following studies in the category of "Studies Currently Under Review" were intended to evaluate Read Naturally and actually implemented the Read Naturally steps:

1. Christ, T. J., & Davie, J. (2009). Empirical evaluation of Read Naturally effects: A Randomized Control Trial (RCT). Unpublished journal article, University of Minnesota, Minneapolis, MN.

This study was reviewed by NCRTI and showed significant effect sizes in fluency and accuracy. Since it was only a ten week study it was not long enough to show an impact on comprehension. As you know the assessments for comprehension are not sensitive enough to pickup comprehension gains in ten weeks.

http://www.readnaturally.com/company/news_seStudyUMn_ncrti.htm

2. Heistad, D. (2008). The effects of Read Naturally on grade 3 reading. Unpublished manuscript.

This study is also included in the NCRTI review and is an appropriate to use for a review of Read Naturally's impact on comprehension since it is a full school year study.

http://www.readnaturally.com/company/news_seStudyUMn_ncrti.htm

3. Hook, P. (2008). Study of Read Naturally. Unpublished study, MGH Institute of Health Professions, Boston, MA.

Pam Hook's email address is phook@mghihp.edu.

Concerning the "Studies Reviewed Under Previous Report", I strongly urge WWC to reevaluate the inclusion of the Hancock and Denton studies as reviews of Read Naturally. The authors of the Hancock, Denton, Kemp and Chenault studies are listed as reviews of Read Naturally. In each case the authors have told WWC that their studies should not be used as an evaluation of Read Naturally because their study was not designed to evaluate Read Naturally and they did not implement the Read Naturally strategy.

I would appreciate it if you could read the following link which includes an analysis of the Hancock and Denton studies as review of Read Naturally.

http://www.readnaturally.com/approach/wwc-summary.htm

Also, under the "Studies Reviewed Under Previous Report" the Heistad and Mesa studies are valid studies of Read Naturally. The Mesa study was designed to evaluate Read Naturally and implemented the Read Naturally strategy. The Mesa study should stand on its own and not be combined with the Hancock study, since the Hancock study did not intend to evaluate Read Naturally and did not implement the Read Naturally strategy.

All the studies under the category "Studies Currently Under Review" were not intended to evaluate Read Naturally and did not implement the Read Naturally strategy except Christ, Heistad, Hook and the control group studies posted on the Read Naturally website.

All the studies under the "Additional Source" section are valid studies of Read Naturally.

I will fax the embargo agreement. I have revised the Intervention Summary. It should be changed on the WWC website as soon as possible. I hope I will be able to talk to you about the serious problems with the current postings of studies as reviews of Read Naturally.

Sincerely,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190

St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:1 Feb 2012 20:57:01 +0000
To:'Info@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot (3477)
Dear Mr. Ihnot,

Thank you for your email. We have forwarded your email and the attachments to the topic area team.

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: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Wednesday, February 01, 2012 11:57 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Dear Susanne James Burdumy,

I appreciate that What Works Clearinghouse (WWC) is updating its Beginning Reading Reviews. It provides an opportunity to correct a serious mistake. The current postings on Read Naturally cause confusion for educators and are a serious injustice to Read Naturally. I would appreciate an opportunity to explain why the Hancock and Denton studies should be removed from WWC website.

The following studies in the category of "Studies Currently Under Review" were intended to evaluate Read Naturally and actually implemented the Read Naturally steps:

1. Christ, T. J., & Davie, J. (2009). *Empirical evaluation of Read Naturally effects: A Randomized Control Trial (RCT).* Unpublished journal article, University of Minnesota, Minneapolis, MN.

This study was reviewed by NCRTI and showed significant effect sizes in fluency and accuracy. Since it was only a ten week study it was not long enough to show an impact on comprehension. As you know the assessments for comprehension are not sensitive enough to pickup comprehension gains in ten weeks.

http://www.readnaturally.com/company/news_seStudyUMn_ncrti.htm

2. Heistad, D. (2008). *The effects of Read Naturally on grade 3 reading.* Unpublished manuscript.

This study is also included in the NCRTI review and is an appropriate to use for a review of Read Naturally's impact on comprehension since it is a full school year study.

http://www.readnaturally.com/company/news_seStudyUMn_ncrti.htm

3. Hook, P. (2008). *Study of Read Naturally*. Unpublished study, MGH Institute of Health Professions, Boston, MA.

Pam Hook's email address is phook@mghihp.edu.

Concerning the "Studies Reviewed Under Previous Report", I strongly urge WWC to reevaluate the inclusion of the Hancock and Denton studies as reviews of Read Naturally. The authors of the Hancock, Denton, Kemp and Chenault studies are listed as reviews of Read Naturally. In each case the authors have told WWC that their studies should not be used as an evaluation of Read Naturally because their study was not designed to evaluate Read Naturally and they did not implement the Read Naturally strategy.

I would appreciate it if you could read the following link which includes an analysis of the Hancock and Denton studies as review of Read Naturally.

http://www.readnaturally.com/approach/wwc-summary.htm

Also, under the "Studies Reviewed Under Previous Report" the Heistad and Mesa studies are valid studies of Read Naturally. The Mesa study was designed to evaluate Read Naturally and implemented the Read Naturally strategy. The Mesa study should stand on its own and not be combined with the Hancock study, since the Hancock study did not intend to evaluate Read Naturally and did not implement the Read Naturally strategy.

All the studies under the category "Studies Currently Under Review" were not intended to evaluate Read Naturally and did not implement the Read Naturally strategy except Christ, Heistad, Hook and the control group studies posted on the Read Naturally website.

All the studies under the "Additional Source" section are valid studies of Read Naturally.

I will fax the embargo agreement. I have revised the Intervention Summary. It should be changed on the WWC website as soon as possible. I hope I will be able to talk to you about the serious problems with the current postings of studies as reviews of Read Naturally.

Sincerely,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally® CELEBRATING 20 YEARS

info@readnaturally.com

From:info@readnaturally.com
Sent:1 Feb 2012 18:41:44 +0000
To:What Works
Subject:Read Naturally - Tom Ihnot
Dear Susanne James Burdumy,

The WWC help desk suggested I email my contact information so that we can discuss the summaries of the Hancock and Denton studies.

Please contact me at 651-286-8721.

I look forward to talking to you.

Thank you,
Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Communications

From:Communications
Sent:2 Feb 2012 13:18:48 -0500
To:What Works
Subject:FW: Attn: Susanne James-Burdumy

From: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Thursday, February 02, 2012 12:59 PM
To: Communications
Subject: Attn: Susanne James-Burdumy

Susanne James-Burdumy,

I appreciate that you are bringing your expertise to reevaluating the What Works Clearinghouse's Beginning Reading section. I have emailed what you requested to WWC. I would appreciate it if you would read the following link which provides an analysis of the Hancock and Denton studies currently listed as reviews of Read Naturally in the Beginning Reading section.

http://www.readnaturally.com/approach/wwc-summary.htm

I think you will see that these studies did not intend to evaluate and did not implement the Read Naturally strategy.

After you have a chance to review the WWC's inclusion of the Hancock and Denton studies, I would appreciate the opportunity to talk to you. My direct line is 651-286-8721.

Thank you,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

info@readnaturally.com

From:info@readnaturally.com
Sent:16 Mar 2012 17:33:57 +0000
To:What Works
Subject:Tom Ihnot - Read Naturally
Susanne James Burdumy,

Christine Tucker has informed me that she will send additional details on her study next week. When I receive the information I will forward it to you.

Also, I would appreciate the opportunity to travel to your office to discuss Read Naturally reviews and your updated review of Beginning Reading.

Please email me a possible date and time.

Thank you,
Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:16 Mar 2012 19:52:49 +0000
To:Info@readnaturally.com
Subject:What Works Clearinghouse (WWCPC 3477)
Dear Mr. Ihnot,

Thank you for contacting the What Works Clearinghouse Help Desk. We will incorporate the information that Christine Tucker provides into our review.

What Works Clearinghouse Help Desk

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/.

WhatWorks

From:WhatWorks
Sent:13 Apr 2012 16:16:00 -0500
To:WhatWorks
Subject:FW: Christine Tucker's article (WWCPC 3477)
Attachments:SCAN1334_000.tif

From: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Friday, April 13, 2012 1:38 PM
To: What Works
Subject: Christine Tucker's article

Dear Susanne James Burdumy,

I have finally obtained a copy of Christine Tucker's article, "Response To Intervention: Increasing Fluency, Rate and Accuracy for Students at Risk for Reading Failure" published in the National Forum of Educational Administration and Supervision Journal, Volume 28, Number 1, 2010-2011.

The article is attached. It is a very positive study of Read Naturally.

Also, I request a thirty minute meeting with you to discuss your review of the Beginning Reading section of the What Works Clearinghouse (WWC) website. I have important information to share with you concerning the use of the Hancock and Denton studies as reviews of Read Naturally.

Please give me a date and time to meet with you at your office. I would greatly appreciate the opportunity.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally® CELEBRATING 20 YEARS

NATIONAL FORUM OF EDUCATIONAL ADMINISTRATION AND SUPERVISION JOURNAL
VOLUME 28, NUMBER 1, 2010-2011

RESPONSE TO INTERVENTION: INCREASING FLUENCY, RATE, AND ACCURACY FOR STUDENTS AT RISK FOR READING FAILURE

Christine Tucker
Walden University

Don Jones
Texas A&M University - Kingsville

ABSTRACT

In recent years, with the advent of NCLB, the U S Department of Education established new guidelines for the identification of students with learning disabilities. In an effort to meet these new guidelines, many districts have implemented a Response to Intervention (RTI) model. This model has been implemented in conjunction with, or in place of the previously common IQ-achievement discrepancy model that emphasized IQ in the identification process. This quasi-experimental study examined the effectiveness of RTI by measurement of the impact of specialized instruction on the reading fluency of identified fourth grade children at risk of failure. The results of the study indicated that student achievement significantly increases when RTI is effectively implemented.

Introduction

The IQ-achievement discrepancy model was developed in 1977 as part of a federally mandated definition of learning disability and is based on the theory that "achievement predicts intelligence, intelligence is a static characteristic, and intelligence serves as a measure of learning capacity" (O'Malley et al., 2002, p. 32). In other words, this model asserts that any child who has been determined to have a low IQ with commensurate achievement is performing up to his or her learning potential, and therefore is not capable of achieving greater success through learning interventions.

must not have been identified as having a specific learning disability in the area of reading, and (d) students must not have been receiving supplemental reading instruction outside of the general education classroom. The decision to use 20 participants was derived based on the relatively small sample size according to current DIBELS Oral Reading Fluency data. The relatively small number of participants allowed the researcher to synthesize and analyze the data for each student in depth to draw sound conclusions about their response to the intervention provided.

Participants' Rights

In order to protect participants' rights, each participant in this study and their legal guardians were provided with a detailed written description of the study including the purpose of the study, details regarding the reading intervention program, potential benefits, and the minimal potential risks. Legal guardians were provided with a letter of consent that they were required to sign if they chose to have their child participate in the study. The form contained contact information regarding where and when the researcher could be reached to field any questions the guardians or the participants may have had regarding the study. Legal guardians and their children's participation were voluntary, and they could have removed their child from the study at any time. Personal information was not used in this study; each student's information was coded to maintain confidentiality, and students' names were changed as well. All student data were stored on the researcher's computer, which required a password to access to protect the participants.

Limitations and Boundaries

This study had limitations that make it difficult to generalize the results to the general population. One important limitation was the relatively small sample size used in this study. The goal of any

research study is to generalize the results from a selected sample back to the entire population (Gravetter & Wallnau, 2005). However, because this study used a sample of only 20 students, additional research corroborating the results would be beneficial in generalizing the results to the population.

Another limitation of this study was that it focused solely on reading fluency and its related components. In order to truly determine if RTI is effective, other areas of learning should be investigated. Moreover, this study was limited to fourth grade students. If this study had used participants from other grade levels the results may have been different. To validate the effects of RTI, interventions should be implemented and studied with students from various grade levels. Also, this study only used one reading intervention. The results may have been different if a variety of reading interventions were used. Consequently, it is necessary to investigate other reading programs to determine if RTI is effective using various interventions.

In addition to the limitations posed by the design of the study, the researcher also could not account for the behavior of the participants in this study. The participants had varying levels of motivation, and therefore they each demonstrated varying amounts of effort and dedication to improving their reading fluency skills. Some students were highly motivated and were extremely focused when participating in the intervention while other students were disinterested and simply went through the motions. Also, it is unknown how much time students in the study practiced reading outside of the allotted time, which may have impacted some of the student's posttest scores.

Data Collection

Following random assignment to either the control group or the quasi-experimental group, each participant was tested (pretest) using form A of the GORT-4 to assess reading fluency, rate, and accuracy. Following the pretest, the students in the control group received

reading instruction using the general education curriculum within the regular education classrooms, and the experimental group received supplementary reading instruction, in addition to the general education reading instruction, using the Read Naturally reading intervention program outside of the regular education setting 4 days per week for 30 minutes across 10 weeks. Following the 10-week period, all participants were retested using form B of the GORT-4 to assess reading fluency (posttest).

Data Analysis

Following the posttest, an independent measures *t* test was used to evaluate the mean difference between the experimental group's growth in the area of fluency, rate, and accuracy based on pretest and posttest scores and the control group's growth in the area of fluency, rate, and accuracy based on pretest and posttest scores. Because the independent measures *t* test only measures the treatment effect, the percentage of variance accounted for by the treatment was also calculated. According to Gravetter and Wallnau (2005), this measure examines whether the treatment causes the scores to vary. In addition, Hartley's F-max test was also used to satisfy the homogeneity of variances assumption, which indicates that the two populations that are being compared have the same variances. Gravetter and Wallnau (2005) indicated that this is an important calculation because if the assumption is not met, the data cannot be interpreted meaningfully from an independent measures experiment.

Fluency

The following hypotheses were tested using an independent measures *t* test.

Ho_1: There is no significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.
Ha_1: There is a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.29 was obtained, which falls within the critical region. Because the *t* score value fell within the critical region, the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest scores of the experimental group is significantly higher than the growth based on pretest and posttest scores of the control group. Figure 1 shows the fluency pretest and posttest scores for each student in the control group, and Figure 2 shows the fluency growth based on pretest and posttest scores for each student in the control group.

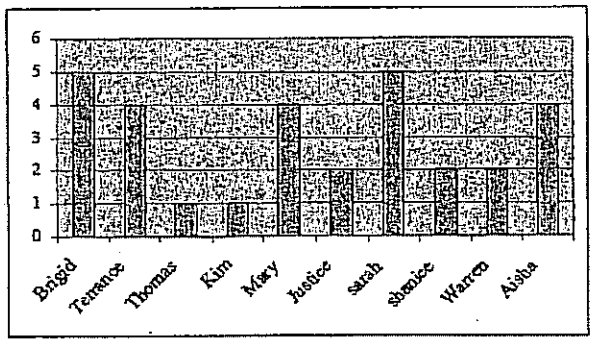


Figure 1: *Experimental group fluency growth*

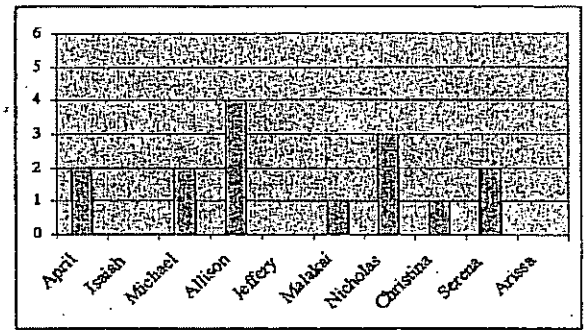


Figure 2: *Control group fluency growth*

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 3.00 with a SD = 1.56 between pretest and posttest scores. The students in the control group had a M = 1.50 with a SD = 1.35 between pretest and posttest scores. This difference was significant, $t(18) = 2.29$, $p < .05$, $r^2 = .23$, which indicates that the 10-week reading intervention using the Read Naturally program had a medium effect on the students' reading fluency scores.

The F-max (1.55) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Rate

The following hypotheses were tested using an independent measures *t* test.

> Ho_1: There is no significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Ha_1: There is a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.51 was obtained, which fell within the critical region. Because the *t* score value fell within the critical region, the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest rate scores of the experimental group is significantly higher than the growth based on pretest and posttest rate scores of the control group. Figure 3 shows the rate pretest and posttest scores for each student in the control group, and Figure 4 shows the rate growth based on pretest and posttest scores for each student in the control group.

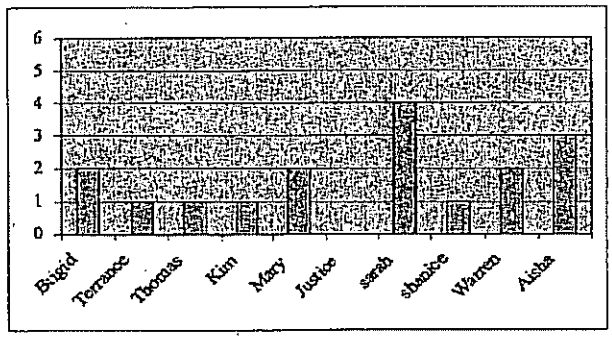


Figure 3: Experimental group rate growth

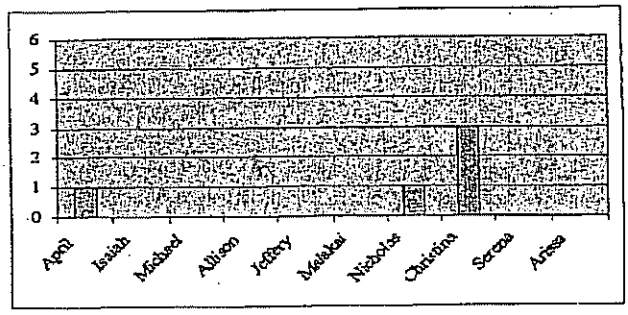


Figure 4: Control group rate growth

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 1.70 with a SD = 1.16 between pretest and posttest scores. The students in the control group had a M = 0.50 with a SD = 0.97 between pretest and posttest scores. This difference was significant, $t(18) = 2.51$, $p < .05$, $r^2 = .26$, which indicates that the 10-week reading intervention using the Read Naturally program had a large effect on students' reading rate scores.

The F-max (1.43) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Accuracy

The following hypotheses were tested using an independent measures *t* test.

> Ho_1: There is no significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Ha_1: There is a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.89 was obtained, which fell within the critical region. Because the *t* score value fell within the critical region the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest rate scores of the experimental group is significantly higher than the growth based on pretest and posttest rate scores of the control group. Figure 5 shows the accuracy pretest and posttest scores for each student in the control group, and Figure 6 shows the accuracy growth based on pretest and posttest scores for each student in the control group.

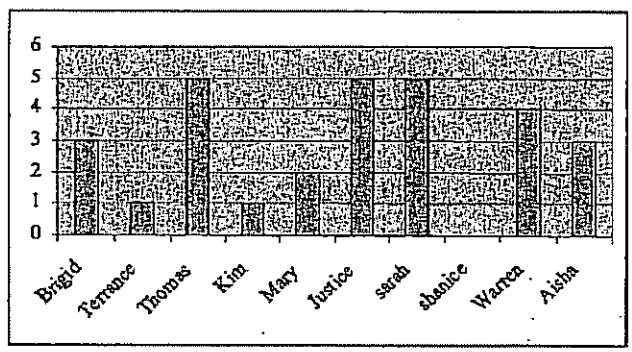


Figure 5: Experimental group accuracy growth

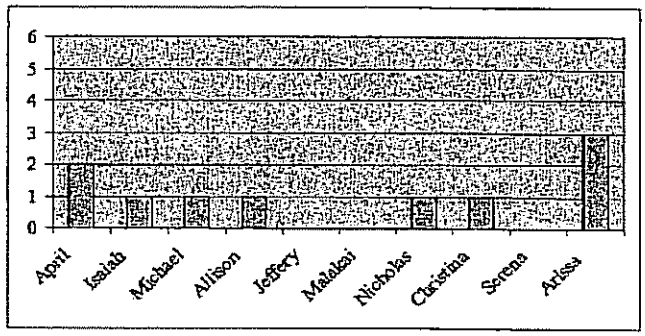


Figure 6: Control group accuracy growth

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 2.90 with a SD = 1.85 between pretest and posttest scores. The students in the control group had a M = 1.00 with a SD = 0.94 between pretest and posttest scores. This information is found on Table 11. This difference was significant, $t(18) = 2.89$, $p < .05$, $r^2 = .32$, which indicates that the 10-week reading intervention using the Read Naturally program had a large effect on the students' reading rate scores.

The F-max (3.49) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Summary

According to the independent measures *t* tests that were conducted, the difference between the experimental group's growth and the control group's growth is significant in all three areas that were examined. Pretest and posttest results indicate that students in the experimental group significantly improved their fluency, rate, and accuracy following the 10-week reading intervention. Although the control group also improved in these areas, the improvement was not

considered to be significant. The null hypotheses, which stated that there is no significant difference in the reading fluency, rate, and accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group, were rejected. Therefore, the answers to the research questions are yes: There is a significant difference in the reading fluency, rate, and accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Recommendations

Implementing the suggestions presented earlier regarding future research could allow the educational system to make sound resolutions on the best way to increase student achievement and identify students with special needs. It will also provide educational systems with the impetus to adopt this method. Administrators in special education departments who determine qualifications for specialized instruction should pay careful attention to research in this area to aid in making decisions regarding the use of RTI. Bergstrom (2008) argued that successfully adopting and implementing RTI goes far beyond progress monitoring and scientifically based interventions; it requires a comprehensive school wide system reform. According to Berkeley, Bender, Peaster, and Saunders (2009), only 15 states have fully developed and are utilizing RTI models. Therefore, the USDOE should look at the results of this study as well as future research to aid in determining the most effective method of identifying learning disabilities, and ultimately implementing new special education standards for all school systems.

The data from this study support RTI as an effective way to improve student achievement. Educators who have students who are at risk for reading failure should also pay attention to the results of this study as well as future research to assist in determining methods of instruction and development of lesson plans and curriculum. The results from this study could be shared with teachers and other faculty

members through professional development to provide educators with a concrete example of how valuable this approach can be with struggling students.

In order for this method to be implemented effectively, educators must develop a deep understanding of and belief in this process. Furthermore, providing teachers and other faculty with the appropriate training will allow them to build confidence in their ability to use RTI with their students. Bergstrom (2008) asserted that in order for RTI to be successful, one of the most important pieces in the reform is professional development geared towards educating teachers about the process. Without the appropriate professional development, the execution of such a program would be impossible.

Conclusion

As stated earlier, the USDOE has created new guidelines for identifying learning disabilities that allow school systems to adopt the RTI method in lieu of or to use in conjunction with the IQ-achievement discrepancy model. Although RTI has been implemented in some school systems throughout the United States, Carney and Stifel (2008) pointed out that "While RTI has come to schools through federal legislation, it has been left to educators and researchers to interpret and investigate the best means of operationalizing this intent to ensure that students difficulties do not stem from instructional deficiencies" (p. 61). In other words, the vague wording of the changes in IDEA leave it open to interpretation. It is up to individual school systems to make decisions on how to structure RTI and ensure students who are at risk are provided with opportunities to become successful. This study, along with future research, will allow the Department of Education, school administrators, and other educational leaders to evaluate the RTI process and establish a protocol for its use in schools across the country to improve student achievement, ultimately effecting social change.

Without relevant research on RTI, many school systems would continue to use the IQ-achievement discrepancy model and perpetuate the ongoing problem of students who do not qualify for special education services falling through the cracks. These students may not be given the chance to increase their achievement skills, even though they struggle with the curriculum simply because they do not qualify to receive specialized instruction (Ukrainetz, 2006). Implementing RTI in schools across the country will allow all students the opportunity to participate in interventions geared towards their specific learning needs. With this targeted instruction, students may be able to improve very precise areas of deficit to advance their overall educational performance, which in turn will afford them greater opportunities in the future.

REFERENCES

Bergstrom, M. K. (2008). Professional development in response to intervention: Implementation of a model in a rural region. *Rural Special Education Quarterly, 27*(4), 27-36.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities, 42*(1), 85-95.

Carney, K. J., & Stiefel, G. S. (2008). Long-term results of problem solving approaches to response to intervention: Discussions and implications. *Learning Disabilities: A Contemporary Journal, 6*(2), 61-75.

Cooter, K. S., & Cooter, R. B. (2004). One size does not fit all: Slow learners in the reading classroom. *The Reading Teacher, 57*(7), 680-684.

Creswell, J. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.

Gravetter, F.J. & Wallnau, L.B. (2005). *Essentials of statistics for the behavioral sciences* (5th ed.). New York, NY: Thomson Publishing.

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34,* 506-522.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities, 38,* 98-108.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disabilities Quarterly, 25*(1), 33-45.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness to intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18*, 157-171.

Hettleman, K. R. (2003). *The invisible dyslexics: How public school systems in Baltimore and elsewhere discriminate against poor children in the diagnosis and treatment of early reading difficulties.* Baltimore, MD: The Abell Foundation.

Johnson, E., Mellard, D. F., & Byrd, S. E. (2005). Alternative models of learning disabilities identification: Considerations and initial conclusions. *Journal of Learning Disabilities, 38*, 569-572.

Kavale, K. A. (2005). Identifying specific learning disability: Is responsiveness to intervention the answer? *Journal of Learning Disabilities, 38*(6), 553-562.

Kavale, K.A., Holdnack, J.A., & Mostert, M.P. (2005). Responsiveness to intervention and the identification of specific learning disability: A critique and alternative proposal. *Learning Disability Quarterly, 28*(1), 2-16.

Lipka, O. Lesaux, N. K., Siegel, & L. S. (2006). Retrospective analyses of the reading development of grade 4 students with reading disabilities: Risk status and profiles over 5 years. *Journal of Learning Disabilities, 39*(5), 364-378.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "'general intelligence,' objectively determined and measured." *Journal of Personality and Social Psychology, 86*(1), 96-111.

Machek, G. R., & Nelson, J. M. (2007). How should reading disabilities be operationalized? A survey of practicing school psychologists. *Learning Disabilities Research and Practice, 22*(2), 147-157.

National Joint Committee on Learning Disabilities (2005). *Response to Intervention (RTI): A primer for parents.* Retrieved from http://www.ldonline.org/article/15857

O'Malley, K. J., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Swank, P. R. (2002). Growth in precursor and reading-related skills: Do low achieving and IQ-discrepant readers develop differently? *Learning Disabilities Research and Practice, 17*, 19-34.

Restori, A. F., Gresham, F. M., & Cook, C. R. (2008). "Old habits die hard:" Past and current issues pertaining to response-to-intervention. *The California School Psychologist, 13*, 67-78.

Ukrainetz, T. A. (2006). The implications of RTI and EBP for SLPs: Commentary on L. M. Justice. *Language, Speech, and Hearing Services in Schools, 37*, 298-303.

U.S. Department of Education. (2008). IDEA Regulations. Retrieved August 19, 2008 from http://idea.ed.gov/explore/view/p/%2Croot%2Cdynamic%2CTopicalBrief%2C23%2C

NATIONAL FORUM JOURNALS

A FAMILY OF JOURNALS ON THE LEADING EDGE OF ACADEMIC EXCELLENCE
FOUNDED 1983

17603 BENDING POST DRIVE, HOUSTON, TX 77095 • (281) 550-5700 • WEBSITE: www.nationalforum.com

October 20, 2010

Dr. Christine Tucker
Richard E. Riley College of Education
Walden University
650 South Exeter Street
Baltimore, Maryland 21202

Dear **Dr. Christine Tucker:**

It is an absolute pleasure for me to contact you. Congratulations on being published in the **NATIONAL FORUM OF EDUCATIONAL ADMINISTRATION AND SUPERVISION JOURNAL, Volume 28, Number 1, 2010.**

Your article earned five affirmative votes from members of our **National Editorial Review Board** and was recommended for national publication by members of our distinguished **National Policy Board.** This journal is a national refereed, juried, peer-review, blind-reviewed periodical. Your manuscript was reviewed by two professors in each of the following states: **Oregon, California, New York, Florida, Michigan,** and **North Carolina.** These referees were selected to review your manuscript by members of our **National Policy Board** representing all **NATIONAL FORUM JOURNALS.** We have about an 88% rejection rate. Competition to publish at the national level remains tense. All **National FORUM Journals** are indexed in **Cabell's Directories.** I invite you and your colleagues to review our world-wide website to see our other refereed journals. Please see: www.nationalforum.com

Over 250,000 guests visit our website yearly. Numerous articles are featured by the nation's foremost thinkers. Over 56,000 articles are downloaded each year for academic purposes.

Founded in 1983, **National FORUM Journals** has published the scholarly contributions of over 4,200 professors with over 2000 articles indexed in **ERIC.** We have fifteen journals in our **National FORUM Family of Journals** that have served scholars nationally and internationally for the past 28 years.

Again, congratulations for publishing in one of our periodicals.

Sincerely,

William Allan Kritsonis

WILLIAM ALLAN KRITSONIS, PHD
EDITOR-IN-CHIEF

William Kritsonis, Ph.D

WAK:am

NATIONAL FORUM JOURNALS

WhatWorks

From:WhatWorks
Sent:13 Apr 2012 21:04:19 +0000
To:Info@readnaturally.com
Subject:RE: Christine Tucker's article
Dear Tom Ihnot,

Thank you for contacting the What Works Clearinghouse (WWC). We have received your email and attachment.

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: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Friday, April 13, 2012 1:38 PM
To: What Works
Subject: Christine Tucker's article

Dear Susanne James Burdumy,

I have finally obtained a copy of Christine Tucker's article, "Response To Intervention: Increasing Fluency, Rate and Accuracy for Students at Risk for Reading Failure" published in the National Forum of Educational Administration and Supervision Journal, Volume 28, Number 1, 2010-2011.

The article is attached. It is a very positive study of Read Naturally.

Also, I request a thirty minute meeting with you to discuss your review of the Beginning Reading section of the What Works Clearinghouse (WWC) website. I have important information to share with you concerning the use of the Hancock and Denton studies as reviews of Read Naturally.

Please give me a date and time to meet with you at your office. I would greatly appreciate the opportunity.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally®
CELEBRATING
20 YEARS

NATIONAL FORUM OF EDUCATIONAL ADMINISTRATION AND SUPERVISION JOURNAL
VOLUME 28, NUMBER 1, 2010-2011

RESPONSE TO INTERVENTION: INCREASING FLUENCY, RATE, AND ACCURACY FOR STUDENTS AT RISK FOR READING FAILURE

Christine Tucker
Walden University

Don Jones
Texas A&M University - Kingsville

ABSTRACT

In recent years, with the advent of NCLB, the U S Department of Education established new guidelines for the identification of students with learning disabilities. In an effort to meet these new guidelines, many districts have implemented a Response to Intervention (RTI) model. This model has been implemented in conjunction with, or in place of the previously common IQ-achievement discrepancy model that emphasized IQ in the identification process. This quasi-experimental study examined the effectiveness of RTI by measurement of the impact of specialized instruction on the reading fluency of identified fourth grade children at risk of failure. The results of the study indicated that student achievement significantly increases when RTI is effectively implemented.

Introduction

The IQ-achievement discrepancy model was developed in 1977 as part of a federally mandated definition of learning disability and is based on the theory that "achievement predicts intelligence, intelligence is a static characteristic, and intelligence serves as a measure of learning capacity" (O'Malley et al., 2002, p. 32). In other words, this model asserts that any child who has been determined to have a low IQ with commensurate achievement is performing up to his or her learning potential, and therefore is not capable of achieving greater success through learning interventions.

must not have been identified as having a specific learning disability in the area of reading, and (d) students must not have been receiving supplemental reading instruction outside of the general education classroom. The decision to use 20 participants was derived based on the relatively small sample size according to current DIBELS Oral Reading Fluency data. The relatively small number of participants allowed the researcher to synthesize and analyze the data for each student in depth to draw sound conclusions about their response to the intervention provided.

Participants' Rights

In order to protect participants' rights, each participant in this study and their legal guardians were provided with a detailed written description of the study including the purpose of the study, details regarding the reading intervention program, potential benefits, and the minimal potential risks. Legal guardians were provided with a letter of consent that they were required to sign if they chose to have their child participate in the study. The form contained contact information regarding where and when the researcher could be reached to field any questions the guardians or the participants may have had regarding the study. Legal guardians and their children's participation were voluntary, and they could have removed their child from the study at any time. Personal information was not used in this study; each student's information was coded to maintain confidentiality, and students' names were changed as well. All student data were stored on the researcher's computer, which required a password to access to protect the participants.

Limitations and Boundaries

This study had limitations that make it difficult to generalize the results to the general population. One important limitation was the relatively small sample size used in this study. The goal of any

research study is to generalize the results from a selected sample back to the entire population (Gravetter & Wallnau, 2005). However, because this study used a sample of only 20 students, additional research corroborating the results would be beneficial in generalizing the results to the population.

Another limitation of this study was that it focused solely on reading fluency and its related components. In order to truly determine if RTI is effective, other areas of learning should be investigated. Moreover, this study was limited to fourth grade students. If this study had used participants from other grade levels the results may have been different. To validate the effects of RTI, interventions should be implemented and studied with students from various grade levels. Also, this study only used one reading intervention. The results may have been different if a variety of reading interventions were used. Consequently, it is necessary to investigate other reading programs to determine if RTI is effective using various interventions.

In addition to the limitations posed by the design of the study, the researcher also could not account for the behavior of the participants in this study. The participants had varying levels of motivation, and therefore they each demonstrated varying amounts of effort and dedication to improving their reading fluency skills. Some students were highly motivated and were extremely focused when participating in the intervention while other students were disinterested and simply went through the motions. Also, it is unknown how much time students in the study practiced reading outside of the allotted time, which may have impacted some of the student's posttest scores.

Data Collection

Following random assignment to either the control group or the quasi-experimental group, each participant was tested (pretest) using form A of the GORT-4 to assess reading fluency, rate, and accuracy. Following the pretest, the students in the control group received

reading instruction using the general education curriculum within the regular education classrooms, and the experimental group received supplementary reading instruction, in addition to the general education reading instruction, using the Read Naturally reading intervention program outside of the regular education setting 4 days per week for 30 minutes across 10 weeks. Following the 10-week period, all participants were retested using form B of the GORT-4 to assess reading fluency (posttest).

Data Analysis

Following the posttest, an independent measures *t* test was used to evaluate the mean difference between the experimental group's growth in the area of fluency, rate, and accuracy based on pretest and posttest scores and the control group's growth in the area of fluency, rate, and accuracy based on pretest and posttest scores. Because the independent measures *t* test only measures the treatment effect, the percentage of variance accounted for by the treatment was also calculated. According to Gravetter and Wallnau (2005), this measure examines whether the treatment causes the scores to vary. In addition, Hartley's F-max test was also used to satisfy the homogeneity of variances assumption, which indicates that the two populations that are being compared have the same variances. Gravetter and Wallnau (2005) indicated that this is an important calculation because if the assumption is not met, the data cannot be interpreted meaningfully from an independent measures experiment.

Fluency

The following hypotheses were tested using an independent measures *t* test.

Ho_1: There is no significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.
Ha_1: There is a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.29 was obtained, which falls within the critical region. Because the *t* score value fell within the critical region, the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest scores of the experimental group is significantly higher than the growth based on pretest and posttest scores of the control group. Figure 1 shows the fluency pretest and posttest scores for each student in the control group, and Figure 2 shows the fluency growth based on pretest and posttest scores for each student in the control group.

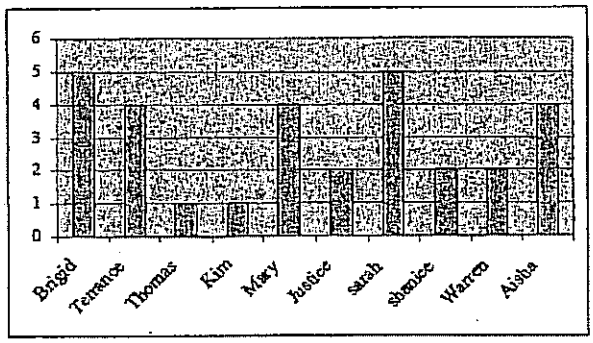


Figure 1: *Experimental group fluency growth*

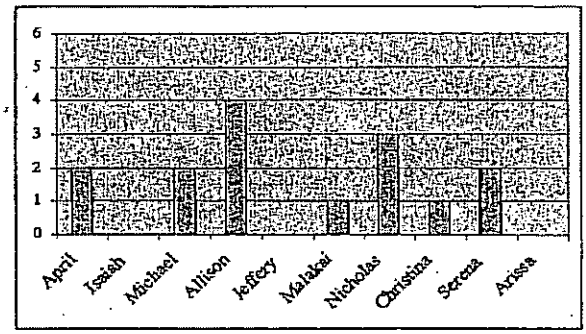


Figure 2: *Control group fluency growth*

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 3.00 with a SD = 1.56 between pretest and posttest scores. The students in the control group had a M = 1.50 with a SD = 1.35 between pretest and posttest scores. This difference was significant, $t(18) = 2.29$, $p < .05$, $r^2 = .23$, which indicates that the 10-week reading intervention using the Read Naturally program had a medium effect on the students' reading fluency scores.

The F-max (1.55) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Rate

The following hypotheses were tested using an independent measures *t* test.

> Ho_1: There is no significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Ha_1: There is a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.51 was obtained, which fell within the critical region. Because the *t* score value fell within the critical region, the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest rate scores of the experimental group is significantly higher than the growth based on pretest and posttest rate scores of the control group. Figure 3 shows the rate pretest and posttest scores for each student in the control group, and Figure 4 shows the rate growth based on pretest and posttest scores for each student in the control group.

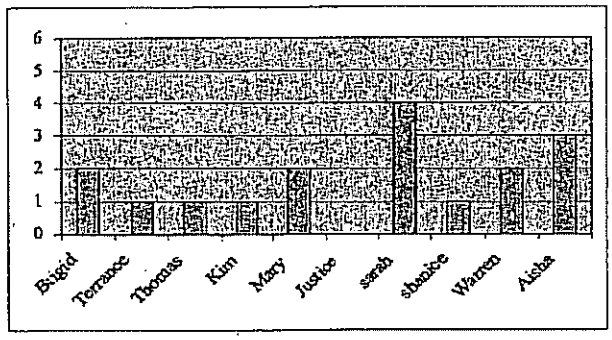


Figure 3: Experimental group rate growth

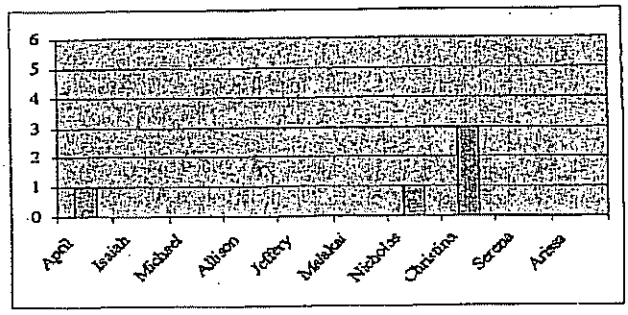


Figure 4: Control group rate growth

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 1.70 with a SD = 1.16 between pretest and posttest scores. The students in the control group had a M = 0.50 with a SD = 0.97 between pretest and posttest scores. This difference was significant, $t(18) = 2.51$, $p < .05$, $r^2 = .26$, which indicates that the 10-week reading intervention using the Read Naturally program had a large effect on students' reading rate scores.

The F-max (1.43) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Accuracy

The following hypotheses were tested using an independent measures *t* test.

> Ho_1: There is no significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Ha_1: There is a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.89 was obtained, which fell within the critical region. Because the *t* score value fell within the critical region the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest rate scores of the experimental group is significantly higher than the growth based on pretest and posttest rate scores of the control group. Figure 5 shows the accuracy pretest and posttest scores for each student in the control group, and Figure 6 shows the accuracy growth based on pretest and posttest scores for each student in the control group.

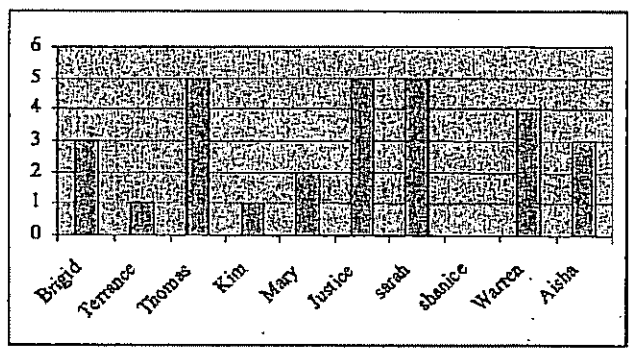


Figure 5: Experimental group accuracy growth

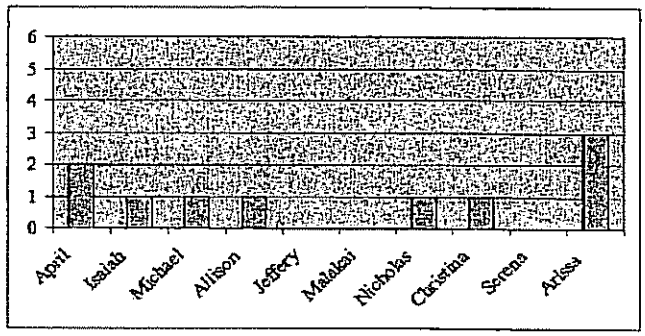


Figure 6: Control group accuracy growth

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 2.90 with a SD = 1.85 between pretest and posttest scores. The students in the control group had a M = 1.00 with a SD = 0.94 between pretest and posttest scores. This information is found on Table 11. This difference was significant, $t(18) = 2.89$, $p < .05$, $r^2 = .32$, which indicates that the 10-week reading intervention using the Read Naturally program had a large effect on the students' reading rate scores.

The F-max (3.49) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Summary

According to the independent measures *t* tests that were conducted, the difference between the experimental group's growth and the control group's growth is significant in all three areas that were examined. Pretest and posttest results indicate that students in the experimental group significantly improved their fluency, rate, and accuracy following the 10-week reading intervention. Although the control group also improved in these areas, the improvement was not

considered to be significant. The null hypotheses, which stated that there is no significant difference in the reading fluency, rate, and accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group, were rejected. Therefore, the answers to the research questions are yes: There is a significant difference in the reading fluency, rate, and accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Recommendations

Implementing the suggestions presented earlier regarding future research could allow the educational system to make sound resolutions on the best way to increase student achievement and identify students with special needs. It will also provide educational systems with the impetus to adopt this method. Administrators in special education departments who determine qualifications for specialized instruction should pay careful attention to research in this area to aid in making decisions regarding the use of RTI. Bergstrom (2008) argued that successfully adopting and implementing RTI goes far beyond progress monitoring and scientifically based interventions; it requires a comprehensive school wide system reform. According to Berkeley, Bender, Peaster, and Saunders (2009), only 15 states have fully developed and are utilizing RTI models. Therefore, the USDOE should look at the results of this study as well as future research to aid in determining the most effective method of identifying learning disabilities, and ultimately implementing new special education standards for all school systems.

The data from this study support RTI as an effective way to improve student achievement. Educators who have students who are at risk for reading failure should also pay attention to the results of this study as well as future research to assist in determining methods of instruction and development of lesson plans and curriculum. The results from this study could be shared with teachers and other faculty

members through professional development to provide educators with a concrete example of how valuable this approach can be with struggling students.

In order for this method to be implemented effectively, educators must develop a deep understanding of and belief in this process. Furthermore, providing teachers and other faculty with the appropriate training will allow them to build confidence in their ability to use RTI with their students. Bergstrom (2008) asserted that in order for RTI to be successful, one of the most important pieces in the reform is professional development geared towards educating teachers about the process. Without the appropriate professional development, the execution of such a program would be impossible.

Conclusion

As stated earlier, the USDOE has created new guidelines for identifying learning disabilities that allow school systems to adopt the RTI method in lieu of or to use in conjunction with the IQ-achievement discrepancy model. Although RTI has been implemented in some school systems throughout the United States, Carney and Stifel (2008) pointed out that "While RTI has come to schools through federal legislation, it has been left to educators and researchers to interpret and investigate the best means of operationalizing this intent to ensure that students difficulties do not stem from instructional deficiencies" (p. 61). In other words, the vague wording of the changes in IDEA leave it open to interpretation. It is up to individual school systems to make decisions on how to structure RTI and ensure students who are at risk are provided with opportunities to become successful. This study, along with future research, will allow the Department of Education, school administrators, and other educational leaders to evaluate the RTI process and establish a protocol for its use in schools across the country to improve student achievement, ultimately effecting social change.

Without relevant research on RTI, many school systems would continue to use the IQ-achievement discrepancy model and perpetuate the ongoing problem of students who do not qualify for special education services falling through the cracks. These students may not be given the chance to increase their achievement skills, even though they struggle with the curriculum simply because they do not qualify to receive specialized instruction (Ukrainetz, 2006). Implementing RTI in schools across the country will allow all students the opportunity to participate in interventions geared towards their specific learning needs. With this targeted instruction, students may be able to improve very precise areas of deficit to advance their overall educational performance, which in turn will afford them greater opportunities in the future.

REFERENCES

Bergstrom, M. K. (2008). Professional development in response to intervention: Implementation of a model in a rural region. *Rural Special Education Quarterly, 27*(4), 27-36.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities, 42*(1), 85-95.

Carney, K. J., & Stiefel, G. S. (2008). Long-term results of problem solving approaches to response to intervention: Discussions and implications. *Learning Disabilities: A Contemporary Journal, 6*(2), 61-75.

Cooter, K. S., & Cooter, R. B. (2004). One size does not fit all: Slow learners in the reading classroom. *The Reading Teacher, 57*(7), 680-684.

Creswell, J. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.

Gravetter, F.J. & Wallnau, L.B. (2005). *Essentials of statistics for the behavioral sciences* (5th ed.). New York, NY: Thomson Publishing.

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34,* 506-522.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities, 38,* 98-108.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disabilities Quarterly, 25*(1), 33-45.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness to intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18*, 157-171.

Hettleman, K. R. (2003). *The invisible dyslexics: How public school systems in Baltimore and elsewhere discriminate against poor children in the diagnosis and treatment of early reading difficulties.* Baltimore, MD: The Abell Foundation.

Johnson, E., Mellard, D. F., & Byrd, S. E. (2005). Alternative models of learning disabilities identification: Considerations and initial conclusions. *Journal of Learning Disabilities, 38*, 569-572.

Kavale, K. A. (2005). Identifying specific learning disability: Is responsiveness to intervention the answer? *Journal of Learning Disabilities, 38*(6), 553-562.

Kavale, K.A., Holdnack, J.A., & Mostert, M.P. (2005). Responsiveness to intervention and the identification of specific learning disability: A critique and alternative proposal. *Learning Disability Quarterly, 28*(1), 2-16.

Lipka, O. Lesaux, N. K., Siegel, & L. S. (2006). Retrospective analyses of the reading development of grade 4 students with reading disabilities: Risk status and profiles over 5 years. *Journal of Learning Disabilities, 39*(5), 364-378.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "'general intelligence,' objectively determined and measured." *Journal of Personality and Social Psychology, 86*(1), 96-111.

Machek, G. R., & Nelson, J. M. (2007). How should reading disabilities be operationalized? A survey of practicing school psychologists. *Learning Disabilities Research and Practice, 22*(2), 147-157.

National Joint Committee on Learning Disabilities (2005). *Response to Intervention (RTI): A primer for parents.* Retrieved from http://www.ldonline.org/article/15857

O'Malley, K. J., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Swank, P. R. (2002). Growth in precursor and reading-related skills: Do low achieving and IQ-discrepant readers develop differently? *Learning Disabilities Research and Practice, 17*, 19-34.

Restori, A. F., Gresham, F. M., & Cook, C. R. (2008). "Old habits die hard:" Past and current issues pertaining to response-to-intervention. *The California School Psychologist, 13*, 67-78.

Ukrainetz, T. A. (2006). The implications of RTI and EBP for SLPs: Commentary on L. M. Justice. *Language, Speech, and Hearing Services in Schools, 37*, 298-303.

U.S. Department of Education. (2008). IDEA Regulations. Retrieved August 19, 2008 from http://idea.ed.gov/explore/view/p/%2Croot%2Cdynamic%2CTopicalBrief%2C23%2C

NATIONAL FORUM JOURNALS

A FAMILY OF JOURNALS ON THE LEADING EDGE OF ACADEMIC EXCELLENCE
FOUNDED 1983

17603 BENDING POST DRIVE, HOUSTON, TX 77095 • (281) 550-5700 • WEBSITE: www.nationalforum.com

October 20, 2010

Dr. Christine Tucker
Richard E. Riley College of Education
Walden University
650 South Exeter Street
Baltimore, Maryland 21202

Dear **Dr. Christine Tucker:**

It is an absolute pleasure for me to contact you. Congratulations on being published in the **NATIONAL FORUM OF EDUCATIONAL ADMINISTRATION AND SUPERVISION JOURNAL, Volume 28, Number 1, 2010.**

Your article earned five affirmative votes from members of our **National Editorial Review Board** and was recommended for national publication by members of our distinguished **National Policy Board.** This journal is a national refereed, juried, peer-review, blind-reviewed periodical. Your manuscript was reviewed by two professors in each of the following states: **Oregon, California, New York, Florida, Michigan,** and **North Carolina.** These referees were selected to review your manuscript by members of our **National Policy Board** representing all **NATIONAL FORUM JOURNALS.** We have about an 88% rejection rate. Competition to publish at the national level remains tense. All **National FORUM Journals** are indexed in **Cabell's Directories.** I invite you and your colleagues to review our world-wide website to see our other refereed journals. Please see: www.nationalforum.com

Over 250,000 guests visit our website yearly. Numerous articles are featured by the nation's foremost thinkers. Over 56,000 articles are downloaded each year for academic purposes.

Founded in 1983, **National FORUM Journals** has published the scholarly contributions of over 4,200 professors with over 2000 articles indexed in **ERIC.** We have fifteen journals in our **National FORUM Family of Journals** that have served scholars nationally and internationally for the past 28 years.

Again, congratulations for publishing in one of our periodicals.

Sincerely,

William Allan Kritsonis

WILLIAM ALLAN KRITSONIS, PHD
EDITOR-IN-CHIEF

William Kritsonis, Ph.D

WAK:am

NATIONAL FORUM JOURNALS

info@readnaturally.com

From:info@readnaturally.com
Sent:16 May 2012 20:18:16 +0000
To:What Works
Subject:Jill Constantine's response to Kevin Feldman
Dear Jill Constantine,

I appreciate your response to Dr. Kevin Feldman's inquiry about WWC's review of Read Naturally. I was happy to read that you are reviewing the Beginning Reading Section of the WWC website and that you are updating the Intervention Report for Read Naturally.

In your response to Dr. Kevin Feldman, you quoted the following WWC policy: "The WWC does not screen based on whether the author explicitly intended the study as an evaluation of an intervention or whether the developer indicates implementation was acceptable, but rather whether the study presents a primary analysis of the effect of an intervention." Thus, WWC concluded that its evaluation of the Hancock study of Read Naturally was appropriate.

However, please allow me to clarify that a Read Naturally intervention is defined as Read Naturally materials used in conjunction with the Read Naturally strategy. The materials alone are not an intervention.

In an email to Becki Herman of WWC, Carrie Hancock stated the following: "While I used Read Naturally materials, I did **NOT** fully implement the Read Naturally strategy and my study was **NOT** intended to evaluate the Read Naturally strategy. Rather, the purpose was to determine the impact of ongoing supplemental fluency practice on second grade students' rates of learning to read." By using the materials and not the strategy, Hancock's study did not analyze Read Naturally as an intervention. If WWC's policy is truly that "the study must present a primary analysis of the effect of an intervention," the Hancock study does not qualify. Carrie Hancock addresses this in her 5/16/12 email. (see below)

WWC has noted on its website that the Hancock study excluded Read Naturally's pre-reading instruction and placement system, but these notations are not sufficient to clarify that the Read Naturally strategy was not used. The pre-reading instruction and placement are critical components of the strategy, but teachers unfamiliar with Read Naturally will not understand or appreciate the impact of these omissions. Additionally, there are several other ways in which the Hancock study deviated from the Read Naturally strategy. These deviations are as follows:

- Hancock omitted the prediction step. This step is crucial for setting a comprehension framework.
- Hancock omitted the key words step. This step is crucial for both comprehension and vocabulary development.
- Hancock omitted the placement step, which ensures that a student is working in the correct level of material.
- Hancock did not set goals based on placement. Setting goals ensures that students are properly challenged.
- Hancock did not use Read Naturally's criteria for individualizing instruction. Hancock set the same goal, the same number of read alongs (3), and the same number of practice readings (5) for each student. Consequently, struggling students who could not achieve

the goal after eight readings of the passage would have felt defeated, and students who could easily achieve the goal without eight readings of the passage would have felt bored.

- Hancock did not adjust levels and goals. Read Naturally's adjustment guidelines are critical to maximize progress.
- Hancock did not use Read Naturally's criteria for passing stories and moving on to new material. Students were allowed to work in new material without mastering expression or comprehension questions.
- Hancock used a group size of 10 to 12 students per one teacher, which is much higher than the ratio Read Naturally requires. Hancock's ratio would have made it impossible for a teacher to assist students in a timely manner for cold and final timings. Consequently, students would have wasted valuable reading time while waiting.
- Hancock's deviations from Read Naturally's placement, initial goal setting, adjusting goals and levels, and progress monitoring guidelines would have destroyed the intrinsic motivation that is a hallmark of the Read Naturally strategy.

Indeed, Carrie Hancock states in her dissertation that, while these critical aspects of the Read Naturally strategy were not part of her study, they should be included in fluency interventions. Based on these numerous and critical differences between Hancock's use of Read Naturally materials and the steps and rationale of the Read Naturally strategy, the Hancock study does not qualify as an analysis of a Read Naturally intervention.

As you update the Intervention Report for Read Naturally, please remove the Hancock study. In addition, please allow the Mesa study to stand alone. The Mesa study was intended to evaluate a Read Naturally intervention and is a proper implementation. I appreciate that you will be evaluating the Christ study, but please keep in mind that comprehension assessments are not sensitive enough to pick up growth in only ten weeks. I also encourage you to consider evaluating the Tucker study, which shows Read Naturally's impact on fourth grade learning disabled students reading at a primary level.

Our mission here at Read Naturally is to do whatever we can to help struggling readers. This is why I have been so persistent in urging you to remove the Hancock study and to evaluate additional studies of Read Naturally. I would appreciate a meeting at your office to discuss WWC posted studies since this critical matter for Read Naturally and all educators who rely on your website. Please email me a date and time.

Thank you,

Tom Ihnot
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: Carrie Hancock [mailto:(b)(6)]
Sent: Wednesday, May 16, 2012 2:33 PM
To: info@readnaturally.com
Subject: Re: from Tom Ihnot

Yes and yes.

Sent from my iPhone

On May 16, 2012, at 12:27 PM, "info@readnaturally.com" <Info@readnaturally.com> wrote:

> Carrie,
>
> Thank you for responding quickly. It helps me respond promptly to Jill Constantine.
>
> Is it fair to say that your use of Read Naturally passages was to analyze the impact of fluency practice with second graders? Also, is it fair to say that you did not use the Read Naturally strategy and as a result the primary analysis of your study was not the effect of the Read Naturally strategy as an intervention?
>
> Sincerely,
>
> Tom Ihnot
>
> Read Naturally, Inc.
> 2945 Lone Oak Drive, Suite 190
> St. Paul, MN 55121
> www.readnaturally.com
> www.oneminutereader.com
> 651-452-4085
> 800-788-4085
> 651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:17 May 2012 20:43:52 +0000
To:Info@readnaturally.com
Subject:RE: Jill Constantine's response to Kevin Feldman (WWC 3562)
Dear Mr. Ihnot,

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

What Works Clearinghouse

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

From: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Wednesday, May 16, 2012 4:18 PM
To: What Works
Subject: Jill Constantine's response to Kevin Feldman

Dear Jill Constantine,

I appreciate your response to Dr. Kevin Feldman's inquiry about WWC's review of Read Naturally. I was happy to read that you are reviewing the Beginning Reading Section of the WWC website and that you are updating the Intervention Report for Read Naturally.

In your response to Dr. Kevin Feldman, you quoted the following WWC policy: "The WWC does not screen based on whether the author explicitly intended the study as an evaluation of an intervention or whether the developer indicates implementation was acceptable, but rather whether the study presents a primary analysis of the effect of an intervention." Thus, WWC concluded that its evaluation of the Hancock study of Read Naturally was appropriate.

However, please allow me to clarify that a Read Naturally intervention is defined as Read Naturally materials used in conjunction with the Read Naturally strategy. The materials alone are not an intervention.

In an email to Becki Herman of WWC, Carrie Hancock stated the following: "While I used Read Naturally materials, I did **NOT** fully implement the Read Naturally strategy and my study was **NOT** intended to evaluate the Read Naturally strategy. Rather, the purpose was to determine the impact of ongoing supplemental fluency practice on second grade students' rates of learning to read." By using the materials and not the strategy, Hancock's study did not analyze Read Naturally as an intervention. If WWC's policy is truly that "the study must present a primary analysis of the effect of an intervention," the Hancock study does not qualify. Carrie Hancock addresses this in her 5/16/12 email. (see below)

WWC has noted on its website that the Hancock study excluded Read Naturally's pre-reading instruction and placement system, but these notations are not sufficient to clarify that the Read

Naturally strategy was not used. The pre-reading instruction and placement are critical components of the strategy, but teachers unfamiliar with Read Naturally will not understand or appreciate the impact of these omissions. Additionally, there are several other ways in which the Hancock study deviated from the Read Naturally strategy. These deviations are as follows:

- Hancock omitted the prediction step. This step is crucial for setting a comprehension framework.
- Hancock omitted the key words step. This step is crucial for both comprehension and vocabulary development.
- Hancock omitted the placement step, which ensures that a student is working in the correct level of material.
- Hancock did not set goals based on placement. Setting goals ensures that students are properly challenged.
- Hancock did not use Read Naturally's criteria for individualizing instruction. Hancock set the same goal, the same number of read alongs (3), and the same number of practice readings (5) for each student. Consequently, struggling students who could not achieve the goal after eight readings of the passage would have felt defeated, and students who could easily achieve the goal without eight readings of the passage would have felt bored.
- Hancock did not adjust levels and goals. Read Naturally's adjustment guidelines are critical to maximize progress.
- Hancock did not use Read Naturally's criteria for passing stories and moving on to new material. Students were allowed to work in new material without mastering expression or comprehension questions.
- Hancock used a group size of 10 to 12 students per one teacher, which is much higher than the ratio Read Naturally requires. Hancock's ratio would have made it impossible for a teacher to assist students in a timely manner for cold and final timings. Consequently, students would have wasted valuable reading time while waiting.
- Hancock's deviations from Read Naturally's placement, initial goal setting, adjusting goals and levels, and progress monitoring guidelines would have destroyed the intrinsic motivation that is a hallmark of the Read Naturally strategy.

Indeed, Carrie Hancock states in her dissertation that, while these critical aspects of the Read Naturally strategy were not part of her study, they should be included in fluency interventions. Based on these numerous and critical differences between Hancock's use of Read Naturally materials and the steps and rationale of the Read Naturally strategy, the Hancock study does not qualify as an analysis of a Read Naturally intervention.

As you update the Intervention Report for Read Naturally, please remove the Hancock study. In addition, please allow the Mesa study to stand alone. The Mesa study was intended to evaluate a Read Naturally intervention and is a proper implementation. I appreciate that you will be evaluating the Christ study, but please keep in mind that comprehension assessments are not sensitive enough to pick up growth in only ten weeks. I also encourage you to consider evaluating the Tucker study, which shows Read Naturally's impact on fourth grade learning disabled students reading at a primary level.

Our mission here at Read Naturally is to do whatever we can to help struggling readers. This is why I have been so persistent in urging you to remove the Hancock study and to evaluate additional studies of Read Naturally. I would appreciate a meeting at your office to discuss WWC posted studies since this critical matter for Read Naturally and all educators who rely on your website. Please email me a date and time.

Thank you,

Tom Ihnot
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

From: Carrie Hancock [mailto:(b)(6)]
Sent: Wednesday, May 16, 2012 2:33 PM
To: info@readnaturally.com
Subject: Re: from Tom Ihnot

Yes and yes.

Sent from my iPhone

On May 16, 2012, at 12:27 PM, "info@readnaturally.com" <Info@readnaturally.com> wrote:

> Carrie,
>
> Thank you for responding quickly. It helps me respond promptly to Jill Constantine.
>
> Is it fair to say that your use of Read Naturally passages was to analyze the impact of fluency practice with second graders? Also, is it fair to say that you did not use the Read Naturally strategy and as a result the primary analysis of your study was not the effect of the Read Naturally strategy as an intervention?
>
> Sincerely,
>
> Tom Ihnot
>
> Read Naturally, Inc.
> 2945 Lone Oak Drive, Suite 190
> St. Paul, MN 55121
> www.readnaturally.com
> www.oneminutereader.com

651-452-4085
800-788-4085
651-452-9204 - fax

February 26, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013

Dear Mr. Ihnot:

Thank you for your inquiries concerning the What Works Clearinghouse (WWC) reviews of four studies of Read Naturally®—Kemp (2006), Denton et al. (2004), Hancock (2002), and Chenault et al. (2006)—and for your email regarding my response to Dr. Kevin Feldman's inquiry about the WWC reviews of studies of Read Naturally®. In response to your concerns, we conducted an independent quality review. The WWC quality review team responds to concerns raised by study authors, curriculum developers, or other relevant parties about WWC reviews published on our website. When a quality review is conducted, a researcher who was not involved in the initial reviews undertakes an independent assessment of the studies in question. The researcher also investigates the procedures used and decisions made during the original reviews of the studies. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the WWC.

The quality review considered the concerns from your emails regarding study inclusion, implementation, student sample, and outcomes.

Study Inclusion

The WWC followed protocol in choosing to review these four studies. The WWC screens studies based on a number of factors, including relevancy and methodology criteria (as described in the WWC Procedures and Standards Handbook, pages 8–10, available at http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19). The WWC does not screen based on whether the author(s) explicitly intended the study as an evaluation of an intervention, but rather whether the study presents a primary analysis of the effect of an intervention. This screening process allows for the review of a wide range of relevant and methodologically sound studies.

Thank you for bringing to my attention the studies by Mesa (2004) and Tucker (2010). The Tucker (2010) study will be included in our literature search for updating the Intervention Report

for Read Naturally®, reviewed under the Beginning Reading Evidence Review Protocol. In addition, as you note, the Mesa (2004) study was included in the original review of Read Naturally®, but will be reviewed again in an updated review.

Implementation

The WWC followed protocol in the manner in which the four studies are described in WWC publications. The WWC Handbook says, "The WWC makes no adjustments or corrections for variations in implementation of the intervention; however, if a study meets standards and is included in an intervention report, descriptions of implementation are provided in the report appendices to provide context for the findings" (page 16). This approach is appropriate because there is no standard metric for fidelity to intervention design. Thus, the WWC includes studies with variation in fidelity and does not evaluate implementation fidelity. The quality review team verified that variations in implementation that are noted in the four studies and that may affect the interpretation of findings were properly included in the WWC publications.

Regarding the Kemp (2006) study, your emails raised a concern about implementation of the Read Naturally® program. The quality review noted that the study does not provide any indication of deviation from program design. Specifically, the study states, "… it could be concluded that all teachers implemented the Read Naturally® program as prescribed" (page 40). Based on this information, the quality review concluded that there was no evidence that deviations in implementation should have been noted in the WWC Intervention Report. Regarding your specific concern about the teacher-to-student ratio, the study does not note any deviation from the Read Naturally® program in this area and does not provide any information on the ratio.

You also expressed concern over deviations in the implementation of Read Naturally® in the Denton et al. (2004) study. The quality review found that the inclusion of additional activities was noted in the study and reported in the WWC Intervention Report (see footnote 7 on page 3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). The other potential deviations that you noted (accountability to pass criteria, use of audio tapes, and the number of sessions per week) were not identified in the study and are therefore not noted in the WWC publication.

Regarding your concerns that the Read Naturally® strategy was not implemented fully for the Hancock (2002) study, the quality review found that WWC procedures were followed in the WWC Intervention Report. The WWC contacted Hancock regarding the implementation of Read Naturally® in the 2002 study. Her response indicated that the study excluded Read Naturally®'s pre-reading vocabulary instruction component and the placement system to individualize instruction. The WWC Intervention Report was revised to note these deviations in implementation (see footnote 4 on page 2, available at http://ies.ed.gov/ncee/wwc/InterventionReport.aspx?sid=407). Other potential concerns with implementation described in your email were not reported by the author and were not included in the revised WWC Intervention Report.

The quality review also investigated concerns that you raised in prior emails. The email from June 2012 regarding Dr. Feldman's questions was addressed in my letter to Dr. Feldman; I will not

repeat the quality review findings here. Your email from May 2012 included some text that was reported to be a quote from Hancock's email to the WWC, but the quoted statements were not part of Hancock's email text. The misreported quote is also found on the Read Naturally® website at http://www.readnaturally.com/approach/hancock.htm.

Regarding the Chenault et al. (2006) study, you expressed concern with the implementation of Read Naturally® related to the selection of students and the limited number of sessions. The quality review found that the WWC Intervention Report appropriately notes both of these factors (see pages 2–3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=409).

Student Sample

You also raised a concern that the Kemp (2006) study included a large share of students who were above the 50th percentile in reading fluency who were not "at-risk" readers. The quality review verified that the WWC Intervention Report accurately states that the study population for the results reported by the WWC was third-grade English language learner (ELL) students (page 2, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). Furthermore, the quality review noted that Read Naturally® program materials state that the program "… products help students of all ages and abilities—whether special education, ELL, Title I, or mainstream—to improve their reading skills, gain confidence, and transform their attitudes toward school and learning" (available at http://www.readnaturally.com/products/default.htm).

Finally, regarding the differences in pre-test performance among native-speaking students in the treatment and control groups, the quality review found that these concerns are not relevant to this WWC publication, which reports on findings only for English language learner students. Furthermore, the quality review concluded that the WWC appropriately restricted the reported findings to impacts assessed relative to the comparison group, and thus did not include achievement gains in oral reading fluency.

Outcomes

Regarding the Denton et al. (2004) study, you expressed concern that the study did not assess fluency. The quality review noted that the WWC Intervention Report clearly states that the study assessed reading achievement (page 4, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408), an outcome of relevance to the WWC ELL topic area. The quality review noted that the Read Naturally® program materials state there are several components of the program beyond fluency that are related to reading achievement: "Our industry-leading programs develop and support the five essential components of reading, identified by the National Reading Panel: phonemic awareness, phonics, fluency, vocabulary, and comprehension" (see "The Read Naturally Strategy," available at http://www.readnaturally.com/approach/default.htm).

I hope that this letter has addressed your concerns. You requested that we meet to discuss the WWC reviews of Read Naturally®. However, WWC policy requires all correspondence in writing. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov.

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

cc: Joy Lesnick, IES

Sent:26 Feb 2013 20:09:42 +0000
To:'kramy@readnaturally.com';'info@readnaturally.com'
Subject:What Works Clearinghouse (QR2010013)
Attachments:QRT2010013.pdf
Dear Mr. Ihnot:

Attached is a response to the questions you raised in your inquiries concerning the What Works Clearinghouse reviews of studies of Read Naturally®.

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/.

info@readnaturally.com

From:info@readnaturally.com
Sent:14 Jun 2012 16:09:54 +0000
To:What Works
Subject:Beginning Reading Reevaluation
Attachments:ProtocolforIntensiveInstructionalIntervention.docx, tucker-jones-article.pdf, Read Naturally Dissertation_Tucker.docx

Beginning Reading Reevaluation,

I have attached the Tucker study of Read Naturally to be evaluated by What Works Clearinghouse (WWC).

Also, since there are serious issues with the four studies WWC has posted as reviews of Read Naturally, I think I should have a chance to meet with a member of the quality review team.

http://www.readnaturally.com/approach/wwc-summary.htm

Please email me a time and place for me to meet with the quality review team to discuss the four studies that are providing inaccurate information to educators.

Thank you,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

NATIONAL CENTER ON INTENSIVE INTERVENTION

PROTOCOL FOR EVALUATING INTENSIVE INSTRUCTIONAL INTERVENTION PROGRAMS

May 2012

Intensive Instructional Intervention Definition
Additional or alternative intervention programs to the core curriculum conducted in small groups or individually with evidence of efficacy for improving academic outcomes for students whose performance is unsatisfactory in the core program

PLEASE READ AND COMPLETE BEFORE YOU START

To determine if your program qualifies for review, please answer the following five questions regarding your program[1]:

(1) Is your program available for dissemination at the current time?	☑ Yes ☐ No
(2) Can you provide evidence* on the effects of your program with students at risk for poor academic outcomes? **By evidence, we mean data from one or more evaluation studies of the program submitted. Studies of similar intervention programs, even if similar to your program, will not be considered as adequate evidence for the purposes of this review.*	☑ Yes ☐ No
(3) Does the evidence come from a published or unpublished study or technical report that may be obtained?	
a. Does the evidence address the effects of the overall program rather than individual components of the program?	☑ Yes ☐ No
(4) Does your intervention meet the following criteria?	
a. Intervention program was delivered in small group or individually.	☑ Yes ☐ No
b. Intervention program occurred over a minimum of 60 minutes a week for 8 weeks.	☑ Yes ☐ No
c. Characteristics and training of the instructors are described in a users' manual or on a website.	☑ Yes ☐ No
d. Intervention program is described in sufficient detail in a users' manual or on a website so that others can use as conducted in the study.	☑ Yes ☐ No
(5) Does the study include the following elements of an evaluation design?	
a. Psychometric properties (e.g., reliability) of the dependent measures are described.	☑ Yes ☐ No
b. Outcome is a quantitative index of students' academic performance.	☑ Yes ☐ No
c. Treatment and control groups are adequately defined.	☑ Yes ☐ No
d. The treatment group is compared to a "business-as-usual" control group.	☑ Yes ☐ No

[1] At this time, the TRC does not consider programs for which only single-subject studies are available to document efficacy.

If you cannot answer yes to all of these questions, we will not review your program. If you are able to answer yes to all of the above questions, then your program qualifies for review. Please proceed to the following page and begin filling out the protocol.

The National Center on Intensive Intervention will review all of your submitted materials to ensure that they adhere to the qualifications for review, stated above. If it is found that your submission packet needs substantial supplemental information or is missing critical information, the entire packet will be returned to you. A revised packet may be re-submitted.

Results of the review will be posted on the Center's website, in the Instructional Intervention Chart. Once the review has begun, withdrawal from the process will not be permitted.

If you are unfamiliar with any of the terms in the protocol or you are unsure of what information to provide, Center staff will be available to assist you. Please contact the National Center on Intensive Intervention:

National Center on Intensive Intervention
American Institutes for Research
1000 Thomas Jefferson Street, NW
Washington, DC 20007

Website: www.intensiveintervention.org
Email Address: NCII@air.org
Toll-free: 866-577-5787

SECTION I: BASIC INFORMATION

A. PROGRAM

Title: Read Naturally

Developer: Candyce Ihnot

Publisher: Read Naturally

Publication Date: 1991

Contact Person: **Name**: Tom Ihnot

Telephone: (651) 452-4085

Email Address: info@readnaturally.com

B. DESCRIPTIVE INFORMATION

Description of program: The Read Naturally strategy develops fluency, supports vocabulary, and promotes comprehension by combining the research-based strategies of teacher modeling, repeated reading, and progress monitoring. A student works on fluency at his or her own pace in an appropriate level of material. The student masters a story by reading along with audio and then practicing the story until he or she can read it accurately and with expression at a goal rate. The student tracks progress on a graph.ck here to enter text.

1. **The program is intended for use in grade(s) (check all that apply):**

☐ Kindergarten	☑ Fifth Grade	☑ Tenth Grade
☑ First Grade	☑ Sixth Grade	☑ Eleventh Grade
☑ Second Grade	☑ Seventh Grade	☑ Twelfth Grade
☑ Third Grade	☑ Eighth Grade	
☑ Fourth Grade	☑ Ninth Grade	

2. **The program is intended for use with (check all that apply):**

- ☐ Students with Disabilities Only
- ☑ Learning Disabilities
- ☐ Mental Retardation
- ☐ Behavioral Disabilities
- ☑ English Language Learners
- ☑ Any Student at Risk of Academic Failure
- ☐ Other: _____

SECTION I: BASIC INFORMATION

3. The academic area of focus is (check all that apply):

- [x] Reading:
  - [ ] Phonological Awareness
  - [x] Phonics / Word Study
  - [x] Comprehension
  - [x] Fluency
  - [x] Vocabulary
  - [ ] Spelling
  - [ ] Other:
- [ ] Math:
  - [ ] Computation
  - [ ] Concepts and/or Word Problems
  - [ ] Whole Number Arithmetic
  - [ ] Comprehensive: Includes computation/procedures, problem solving and mathematical concepts
  - [ ] Algebra
  - [ ] Fractions, Decimals (Rational Number)
  - [ ] Geometry and Measurement
  - [ ] Other:
- [ ] Writing
  - [ ] Handwriting
  - [ ] Spelling
  - [ ] Sentence Construction
  - [ ] Planning and Revising
  - [ ] Other:

4. Acquisition and Cost Information:

Where to Obtain: Read Naturally, Inc.

Address: 2945 Lone Oak Dr, Suite 190, St. Paul, MN 55121

Phone #: (651) 452-4085

Web Site: www.readnaturally.com

Initial cost per student for implementing program: Encore $129 per level; 30 seats of Read Live $599; Network Plus Software levels $399 each

Replacement cost per student for subsequent use: Read Live is a yearly subscription; Encore and Network Plus software one time purchase.

Additional Cost Information: Describe basic pricing plan and/or structure of the program. Also, provide information on what is included in the published program, as well as what is not included but required for implementation: Click here to enter text.

5. **Extent of Program Use by Practitioners: Describe where the program has been used and by whom (e.g., numbers of districts/schools, locations, years in use) and how you know this**: Read Naturally is used by thousands of school in all fifty states and internationally.

C. PROGRAM SPECIFICATIONS

1. **Is the program designed for use with (check all that apply)?**

   ☑ Individual Students

   ☑ Small Groups of Students

2. **If group-delivered, how many students comprise a small group**? The Read Naturally strategy individualizes instruction within a group setting. A teacher/student ratio of one to six is recommended.

3. **Program administration time**:

   30 Minutes per session

   3 Sessions per week

   _____ Number of weeks

4. **Does the program include highly specified teacher manuals?**

   ☑ Yes

   ☐ No

5. **Does the program require technology?**

   ☑ Yes

   ☐ No

   If so, please describe required technology and the extent to which it is combined with teacher small group instruction: The Encore version requires a CD player. The software version is standalone software CD or a school network version. Read Live is an online web based version.

D. TRAINING

1. **Is training for the instructor required?**

   ☑ Yes

   ☐ No

2. **Time required for training instructor:**

   ☐ Training Not Required

   ☐ Less Than 1 Hour of Training

- ☐ 1-4 Hours of Training
- ☑ 4-8 Hours of Training
- ☐ More Than 8 hours of Training. Please specify:
- ☐ Not available

3. **Please describe the format and content of the instructor training:** Read Naturally provides live full day seminars, video workshops and self study training books.

4. **Minimum qualifications of the instructor:**
   - ☑ Professional
   - ☐ Paraprofessional
   - ☐ Other: Please describe:

5. **Does the program assume that the instructor has expertise in a given area?**
   - ☐ No
   - ☑ Yes: Please describe: Reading

6. **Are training manuals and materials available?**
   - ☑ Yes
   - ☐ No

7. **Describe how the training manuals/materials were field-tested with the target population of instructors and students:** Over the past 21 years, the manuals have been updated and revised numerous times based on teacher input.

8. **May practitioners obtain ongoing professional/technical support?**
   - ☐ No
   - ☑ Yes Please specify: Professional educators and technical support staff are available five days a week to support teachers and technical staff.

SECTION II: TECHNICAL INFORMATION (STUDY FORM)

Instructions to Vendors:
Please use this section to provide evidence supporting the efficacy of your program. Evidence must be documented in a published study or technical report. For each study submitted, complete one study form (below) and attach the corresponding publication or technical report.

You may submit a minimum of one and a maximum of 5 study forms during a given review cycle. Copy additional study forms as necessary.

SUMMARY OF EVIDENCE BASE

Please identify, to the best of your knowledge, all of the research studies that have been conducted to date supporting the efficacy of your program. Please provide *citations only*; do not include any descriptive information on these studies. NCII staff will also conduct a search to confirm that the list you provide is accurate.

Christ, T. J. & Davie, J. (2009). Empirical evaluation of Read Naturally effects: A randomized control trial (RCT). (unpublished manuscript). University of Minnesota, Minneapolis, MN.

Hasbrouk, J. E., Ihnot, C., & Rogers, G. H. (1999). Read Naturally: A strategy to increase oral reading fluency. *Reading Research and Instruction (39), 1,* 27-38.

Heistad, D. (2004). The effects of Read Naturally on fluency and reading comprehension: A supplemental service study (four-school study). (unpublished manuscript). Minneapolis, MN.

Heistad, D. (2004). The effects of Read Naturally on fluency and reading comprehension: A supplemental service study (two-school study). (unpublished manuscript). Minneapolis, MN.

Heistad, D. (2004). The effects of Read Naturally on grade 3 reading: A study in the Minneapolis public schools. (unpublished manuscript). Minneapolis, MN.

In the remainder of this section, you may select up to five studies from this evidence base for the TRC to review. We recommend that you select studies that are the most recent and that are the most rigorous in terms of their design. The TRC will review and publish information on the studies you select. However, the TRC also reserves the right to review and rate studies not included in this section, if we feel that important information about the efficacy of your program is missing.

STUDY (please attach)

Do not submit multiple study forms that report data on the same sample. If there are multiple reports on the same sample, complete one study form and cite the multiple publications or reports.

Title: Response to intervention: Increasing fluency, rate, and accuracy for students at risk for reading failure.

Authors: Christine Tucker & Don Jones

Year Published: 2010

Study Citation: Tucker, C. & Jones, D. (2010). Response to intervention: Increasing fluency, rate, and accuracy for students at risk for reading failure. *National Forum of Educational Administration and Supervision Journal, 28(1),* 28-47.

If unpublished, how can user obtain technical report(s): Click here to enter text.

SECTION II: TECHNICAL INFORMATION (STUDY FORM)

Does the Study Form include information not described in the attached report(s)?

☐ Yes

☑ No

If yes, bold this information in the Study Form, and indicate how the user can obtain this additional information:

Click here to enter text.

A. PARTICIPANTS

1. **Describe how students were selected to participate in the study**:
   Students were selected to participate if they received a score of 70 or below on the 4th grade oral reading fluency probe (placing them within the at-risk category on DIBELs probes). Students who participated did not receive special education services and did not receive prior supplemental reading instruction outside of the general education curriculum.

2. **Describe how students were identified as at risk for academic failure**:
   Students were identified as at risk for reading failure based on their performance on DIBELS oral reading fluency probes, scoring 70 or below on a grade level probe.

   a. **Clarify which treatment is the submitted program**:
   Read Naturally is the submitted program.

   b. **Specify which condition is the control condition if there are more than two treatment conditions**:
   Click here to enter text.

   c. **If you have a third, competing condition, in addition to your control and intervention condition, identify what the competing condition is [data from this competing condition will not be used])**:
   Click here to enter text.

3. **Please provide the sample sizes for your study, for all types of participants (schools, instructors, classrooms, and students) and relevant conditions (intervention and control). Alternatively, describe where this information can be found in the attached report:**
   Click here to enter text.

4. **How many program students were pretested?** 10 **How many were posttested?** 10

5. **How many control students were pretested?** 10 **How many were posttested?** 10

6. **What was randomly assigned?**

   ☐ Schools

   ☐ Teachers

☑ Students

7. **What unit(s) were used for primary data analysis?**

☐ Schools

☐ Teachers

☑ Students

B. DESIGN

1. **Was random assignment used?**

☑ Yes

Please provide the location (highlight, page number, etc.) in the attached document that describes the random assignment process: This was a quasi-experimental design. Once students were selected for participation, they were randomly assigned to either the control or treatment group. This information can be found on page 32 of the journal article.

☐ No

Please describe the study design: Click here to enter text.

2. **Using the tables below, provide data demonstrating comparability of the program group and control group on pretest performance and demographics.**

For pretest performance, include (a) pretreatment means and standard deviations on each variable for the program and control groups, (b) sample sizes for the program and control groups.

Pretest academic performance measures (provide for subset of students who completed the study)

Measures (Name)	Program (n=)		Control (n=)		Effect Size: Mean difference (in SD units) *
	Mean	Standard Deviation	Mean	Standard Deviation	
Gray Oral Reading Test-4th, Form A: Rate	7.1	2.04	7.6	1.67	
Gray Oral Reading Test-4th, Form A: Accuracy	8.9	1.55	8.9	2.51	
Gray Oral Reading Test-4th, Form A: Fluency	7.6	1.65	7.7	2.15	

*NCII staff will complete this column.

Demographic Information

	Program		Control	
	Number	Percentage	Number	Percentage
Grade Level				

SECTION II: TECHNICAL INFORMATION (STUDY FORM)

	Program		Control	
	Number	Percentage	Number	Percentage
Kindergarten				
Grade 1				
Grade 2				
Grade 3				
Grade 4	10	100	10	100
Grade 5				
Grade 6				
Grade 7				
Grade 8				
Grade 9				
Grade 10				
Grade 11				
Grade 12				
Race-ethnicity				
African-American	2	20	4	40
American Indian				
Asian/Pacific Islander				
Hispanic				
White	6	60	5	50
Other	2	20	1	10
Socioeconomic Status				
Subsidized Lunch	4	40	3	30
No subsidized Lunch	6	60	7	70
Disability Status				
Speech-Language Impairments				
Learning Disabilities				
Behavior Disorders				
Mental Retardation				
Other				
Not Identified with a Disability				
ELL status				
English Language Learner				
Not English Language Learner				
Gender				
Female	5	50	7	70
Male	5	50	3	30

C. FIDELITY OF IMPLEMENTATION

1. **How was the program delivered (check all that apply)?**

☑ Individually ☐ Small group

Average group size: _____

Range: _____

SECTION II: TECHNICAL INFORMATION (STUDY FORM)

2. **What was the duration of the intervention?**

   Weeks: 10

   Sessions per week: 4

   Duration of sessions in minutes: 30

3. **What were the background, experience, training and ongoing support of the instructors?**
   The investigator/instructor was a speech-language pathologist employed at the school. She was completing her dissertation study to complete her Ed.D degree. The investigator/instructor attended a "Read Naturally" Seminar in 2009 to receive a certificate of completion. She had access to the training manual prior to program implementation. The manual was used throughout the study as necessary.

4. **Describe when and how fidelity of treatment information was obtained.**
   NA

5. **What were the results on the fidelity of treatment implementation measure?**
   NA

6. **Was the fidelity measure also used in control classrooms?**
   NA

SECTION II: TECHNICAL INFORMATION (STUDY FORM)

D. MEASURES

NOTE: Answers to the questions below must be provided separately for *targeted measures* and *broader measures*. The placement of measures into these categories is your responsibility; however, the TRC reserves the right to question and re-categorize measures as necessary, depending on the information provided. Definitions of these types of measures are:

- ***Targeted measures*** assess aspects of competence the program was directly targeted to improve. Typically, this does not mean the very items taught but rather novel items structured similarly to the content addressed in the program. For example, if a program taught word attack, a targeted measure would be decoding of pseudowords. If a program taught comprehension of cause-effect passages, a targeted measure would be answering questions about cause-effect passages structured similarly to those used during intervention, but not including the very passages used for intervention.

- ***Broader measures*** assess aspects of competence that are broader than the skills targeted by the program but not directly taught in the program. For example, if a program taught word-level reading skill, a broader measure would be answering questions about passages the student reads. If a program taught calculation skill, a broader measure would be solving word problems that require the same kinds of calculation skill taught in the program.

Measures related to program implementation or process should NOT be included.

1. **Targeted outcome measures**
   Using the table below, describe each targeted outcome measure in terms of its reliability and relevance to your program's instructional content. Please be specific regarding the type of reliability reported, and regarding similarities between the measure and the content taught by the program. Please also describe the extent to which students in the control group receive instruction related to the outcome measure.

   Please note: Vendors must supply reliability coefficients even for well known measures.

Measure name	**Reliability statistics** (specify type of reliability, e.g. Cronbach's alpha, IRT reliability, temporal stability, inter-rater)	**Relevance to program instructional content**	**Exposure to related content among control group**
Gray Oral Reading Test, 4th Edition	Test-retest, alternate form, Cronbach's alpha: all above .90	The GORT-4 is a measure of fluency and comprehension. The Read Naturally Intervention program targets fluency.	Students in the control group received instruction in the Scott Foresman Reading Street curriculum. This may include elements of fluency instruction, but it was used for all general education students (e.g. "business as usual").

2. Broader outcome measures

Using the table below, describe each broader outcome measure in terms of its reliability and relevance to your program's instructional content. Please be specific regarding the type of reliability reported, and regarding similarities between the measure and the content taught by the program. Please also describe the extent to which students in the control group receive instruction related to the outcome measure.

Please note: Vendors must supply reliability coefficients even for well known measures.

Measure name	**Reliability statistics** (specify type of reliability, e.g. Cronbach's alpha, IRT reliability, temporal stability, inter-rater)	**Relevance to program instructional content**	**Exposure to related content among control group**

If you have excluded a variable or data that are reported in the attached document, explain rationale for exclusion: Click here to enter text.

E. RESULTS

1. **Describe the analyses used.**
   An independent samples t-test was used to determine differences between groups after the intervention ended. Assumptions of homogeneity of variance were satisfied. See page 35 of the journal article for results.

2. **In the tables that follow, for each outcome measure, provide the following posttreatment data: unadjusted mean, adjusted mean (e.g., corrected for pretest differences)[2], unadjusted standard deviation, and sample size. Do this for the program group and for the control group. Data must be included for *each* outcome measure (targeted and broader) that was described in Section D: Measures.**

   NCII staff will compute effect sizes using two standard formulas: (1) effect size based on adjusted posttest means and post test SDs; and (2) effect size based on unadjusted posttest means. The second formula is typically used only in instances in which we can assume pretest group equivalency. Therefore, NCII will only be reporting effect size based on unadjusted posttests for studies that either (a) use random assignment or (b) are quasi-experiments but pretest differences on outcome measures are not statistically significant and fall within 0.25 standard deviations. ***The TRC strongly encourages vendors to submit adjusted means, so that the NCII can report effect sizes for your study using the most rigorous methods.*** Note also that the NCII will not be able to report effect size for any variable for which only posttest data are known, due to the need for pretests in calculating adjusted posttest scores.

[2] For guidance on how to compute adjusted means, see the document "Adjusted posttest means."

SECTION II: TECHNICAL INFORMATION (STUDY FORM)

If available, please also provide posttreatment data disaggregated for:

(1) A sub-sample of students at or below the 20th percentile on pretest measures of academic achievement

(2) Sub-samples of students in different demographic groups, e.g. ELLs, SWDs, or students in specific racial-ethnic categories.

Please copy or use additional tables as necessary to include the additional, disaggregated outcome data.

Outcome data for program and control groups

Check one:

☑ Full sample

☐ Sub-sample of low-achieving students, falling at or below the 20th percentile on pretest measures of achievement (please specify):

Click here to enter text.

☐ Disaggregated data (describe demographic group):

Click here to enter text.

Measure (Name)	**Posttreatment data** Provide for subset of students who completed the study.		
		Program	Control
Gray Oral Reading Test, 4th Edition: Rate	Unadjusted Mean	8.8	8.1
	Adjusted Mean	8.9	7.9
	Unadjusted Standard Deviation	2.04	1.67
	n	10	10
Gray Oral Reading Test, 4th Edition: Accuracy	Unadjusted Mean	11.8	9.9
	Adjusted Mean	11.8	9.9
	Unadjusted Standard Deviation	1.55	2.51
	n	10	10
Gray Oral Reading Test, 4th Edition: Fluency	Unadjusted Mean	10.6	9.2
	Adjusted Mean	10.6	9.1
	Unadjusted Standard Deviation	1.65	2.15
	n	10	10
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		

SECTION II: TECHNICAL INFORMATION (STUDY FORM)

Measure (Name)	**Posttreatment data** Provide for subset of students who completed the study.		
		Program	Control
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		

SECTION II: TECHNICAL INFORMATION (STUDY FORM)

Outcome data for program and control groups

Check one:

☐ Full sample

☐ Sub-sample of low-achieving students, falling at or below the 20th percentile on pretest measures of achievement (please specify):

Click here to enter text.

☐ Disaggregated data (describe demographic group):

Click here to enter text.

Measure (Name)	**Posttreatment data** Provide for subset of students who completed the study.		
		Program	Control
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		

SECTION II: TECHNICAL INFORMATION (STUDY FORM)

Outcome data for program and control groups

Check one:

☐ Full sample

☐ Sub-sample of low-achieving students, falling at or below the 20th percentile on pretest measures of achievement (please specify):

Click here to enter text.

☐ Disaggregated data (describe demographic group):

Click here to enter text.

Measure (Name)	**Posttreatment data** Provide for subset of students who completed the study.		
		Program	Control
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		
	Unadjusted Mean		
	Adjusted Mean		
	Unadjusted Standard Deviation		
	n		

Walden University

COLLEGE OF EDUCATION

This is to certify that the doctoral study by

Christine N. Tucker

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Don Jones, Committee Chairperson, Education Faculty
Dr. Casey Reason, Committee Member, Education Faculty

Chief Academic Officer

Denise DeZolt, Ph.D.

Walden University
2010

Response to Intervention: Increasing Fluency, Rate, and Accuracy for Students At Risk for Reading Failure

By

Christine N. Tucker

M.S., Northeastern University, 2003
B.S. Northeastern University, 2003

Doctoral Study Submitted in Fulfillment
of the Requirements for the Degree of
Doctor of Education

Walden University
June 2010

Abstract

Many schools in the United States use the IQ-achievement discrepancy method to identify children with learning disabilities (a significant split between their intelligence and their achievement skills). Unfortunately struggling students who are not identified by this method remain in the regular education setting with no additional supports and can experience failure throughout their academic careers. In 2004, the Individuals with Disabilities Education Act was reauthorized to provide states with the option of using a response to intervention (RTI) model in lieu of or in conjunction with the IQ-achievement discrepancy model. However, many states have chosen not to adopt this new model due to the lack of research supporting RTI. The purpose of this quasi-experimental pre/posttest design study was to determine the impact of specialized instruction on reading fluency, rate, and accuracy with students in grade 4 who have been identified as at risk for reading failure. Twenty students identified as at risk for reading failure were split into two groups: a control group that received reading instruction within the classroom using the general education curriculum and an experimental group that received supplemental reading instruction (RTI) using Read Naturally in addition to the general curriculum. The difference between pretest and posttest scores obtained from the Gray Oral Reading Test were analyzed using an independent samples *t* test, which revealed that students who participated in the reading intervention demonstrated significantly more growth in the areas of reading fluency, rate, and accuracy than students in the control group. This targeted instruction, using the RTI model, enabled students to increase their overall reading achievement. This research could help educational leaders to promote social change by coming to a consensus on how to best solve this longstanding problem so that we may move forward and continue to make advancements in the field of education, ultimately advancing our society.

Response to Intervention: Increasing Fluency, Rate, and Accuracy for Students At Risk for Reading Failure

By

Christine N. Tucker

M.S., Northeastern University, 2003
B.S. Northeastern University, 2003

Doctoral Study Submitted in Fulfillment
of the Requirements for the Degree of
Doctor of Education

Walden University
June 2010

Acknowledgments

I would like to take this opportunity to thank the many people in my life who have helped me through this process. First, I would like to thank my wonderful, supportive husband who encouraged me every step of the way. Without him, I would not have found the confidence I needed to even begin this process, and I am so grateful that he believed in my ability to accomplish this seemingly impossible task. Without his support I would not have succeeded.

I would also like to thank my caring, empathetic, and encouraging mother who has taught me that with dedication and hard work, anything is possible. Since I was a child, she has supported me and encouraged me to live up to my potential. I never dreamed I had this much potential, but she always knew I did. I am also very grateful for the support that my sister and the rest of my family have given me along the way. I don't know where I would be without their love and guidance.

I could not have accomplished this task without the help of my doctoral chair, Dr. Jones. He has been so patient, understanding, and dedicated to helping me succeed since the very beginning. Without his advice and words of encouragement, I would not have made it this far. I am so thankful for all of the help he has given me throughout this process. I am also grateful to Dr. Reason, my methodologist, for assisting me during this process.

Last but certainly not least, my colleagues have helped me tremendously and I am so thankful to work with such kind, caring, and dedicated people. Their suggestions and willingness to provide me with assistance has been paramount to my success.

Table of Contents

List of Tables iii
List of Figures iv

Section 1: Introduction to the Study 1
Background of Problem 1
Problem Statement 2
Theoretical Framework 4
Purpose of Study 5
Significance 5
Nature of the Study 6
Research Questions 7
Hypotheses 7
Definition of Terms 8
Delimitations, Limitations, and Assumptions 9
Summary 9

Section 2: Literature Review 11
Introduction 11
Historical Background 12
Theoretical Connections 21
Intelligence 21
Response to Intervention 33
Reading Disabilities 47
Conclusion 62

Section 3: Methodology 64
Introduction 64
Specific Research Questions 64
Hypotheses 65
Setting and Sample 65
Study Process and Data Collection 66
Intervention 67
Read Naturally (Software Edition) 67
Instrumentation and Materials 69
DIBELS 69
Gray Oral Reading Test: Fourth Edition 69
Reliability and Validity 70
Read Naturally 70
DIBELS 71
Gray Oral Reading Test: Fourth Edition 71
Data Collection and Analysis 72
Participants' Rights 73
Role of Researcher 73
Measures for Ethical Protection 73

Section 4: Results 74
Introduction 74
Overview of the Study 74
Sample Description 75
Control Group 76
Experimental Group 76
Fluency Pretest and Posttest Results 76
Hypotheses 76
Experimental Group 77
Control Group 78
Statistical Analysis of Fluency Data 80
Independent Measures t test 80
Results 81
Rate Pretest and Posttest Results 82
Hypotheses 82
Experimental Group 82
Control Group 83
Statistical Analysis of Rate Data 85
Independent Measures t test 85
Results 86
Accuracy Pretest and Posttest Results 87
Hypotheses 87
Experimental Group 87
Control Group 88
Statistical Analysis of Accuracy Data 90
Independent Measures t test 90
Results 91
Summary 92

Section 5: Summary, Conclusions, and Recommendations 93
Introduction 93
Overview of the Study 93
Summary of Research Findings 94
Fluency 95
Rate 96
Accuracy 97
Limitations 99
Future Research 100
Recommendations for Action 100
Implications for Social Change 102

REFERENCES 104
Appendix A: Permission to use DIBELS 112
Appendix B: Permission to use Read Naturally 113
Appendix C: Assent Form 114
Appendix D: Consent Form 116

List of Tables

Table 1. Experimental Group Pretest and Posttest Fluency Scores 77
Table 2. Control Group Pretest and Posttest Fluency Scores 79
Table 3. Group Fluency Statistics 82
Table 4. Independent Measures *t* Test Fluency Data 82
Table 5. Experimental Group Pretest and Posttest Rate Scores 83
Table 6. Control Group Pretest and Posttest Rate Scores 84
Table 7. Group Rate Statistics 86
Table 8. Independent Measures *t* Test Rate Data 87
Table 9. Experimental Group Pretest and Posttest Accuracy Scores 88
Table 10. Control Group Pretest and Posttest Accuracy Scores 89
Table 11. Group Accuracy Statistics 91
Table 12. Independent Measures *t* Test Accuracy Data 92

List of Figures

Figure 1. Experimental group pretest and posttest fluency scores 78
Figure 2. Control group pretest and posttest Fluency scores 79
Figure 3. Experimental group fluency growth 80
Figure 4. Control group fluency growth 80
Figure 5. Experimental group pretest and posttest rate scores 83
Figure 6. Control group pretest and posttest rate scores 84
Figure 7. Experimental group rate growth 85
Figure 8. Control group rate growth 85
Figure 9. Experimental group pretest and posttest accuracy scores 88
Figure 10. Control group pretest and posttest accuracy scores 89
Figure 11. Experimental group accuracy growth 90
Figure 12. Control group accuracy growth 90

Section 1: Introduction to the Study

Background of Problem

The IQ-achievement discrepancy model was developed in 1977 as part of a federally mandated definition of learning disability and is based on the theory that "achievement predicts intelligence, intelligence is a static characteristic, and intelligence serves as a measure of learning capacity" (O'Malley et al., 2002, p. 32). In other words, this model asserts that any child who has been determined to have a low IQ with commensurate achievement is performing up to his or her learning potential, and therefore is not capable of achieving greater success through learning interventions. Consequently, using this model has caused many students who are at risk for failure to be denied services to increase achievement (Restori, Gresham, & Cook, 2008).

The IQ-achievement discrepancy model has been used to identify learning disabilities (LD) throughout the United States since that time and in recent years has become a controversial issue in education (Machek & Nelson, 2007). Many researchers have argued that IQ is the most critical aspect of identifying children with LD (Johnson, Mellard, & Byrd, 2005; Kavale, 2005) while others have suggested that IQ is just a small piece of the puzzle, and more focus should be placed on dynamic assessments and responsiveness to intervention (Cooter & Cooter, 2004; Fletcher, Francis, Morris, & Lyon, 2004; Hettleman, 2003; National Joint Committee on Learning Disabilities, 2005; Stuebing, Lyon, Shaywitz, & Shaywitz, 2005; Fuschet al. 2003).

Because of this conflict, many researchers have investigated different models of identifying LD, which spurred the U.S. Department of Education (USDOE) to create new guidelines in the identification process. In 2004, President George W. Bush signed a reauthorization of the Individuals with Disabilities Education Act (IDEA) that stated:

> A state must adopt, consistent with 34 CFR 300.309, criteria for determining whether a child has a specific learning disability as defined in 34 CFR 300.8(c)(10). In addition, the

> criteria adopted by the State: 1) Must not require the use of a severe discrepancy between intellectual ability and achievement for determining whether a child has a specific learning disability, as defined in 34 CFR 300.8(c)(10); 2) Must permit the use of a process based on the child's response to scientific, research-based intervention; and 3) May permit the use of other alternative research-based procedures for determining whether a child has a specific learning disability, as defined in 34 CFR 300.8(c)(10). (U.S. Department of Education, 2008, ¶ 2)

After this reauthorization, school systems across the country began implementing a response to intervention (RTI) model to use in lieu of or in conjunction with the IQ-achievement discrepancy model. However, because this newer method of identification is still optional, many school systems continue to use the discrepancy model (Restori, Gresham, & Cook, 2008).

Problem Statement

There is a problem in the United States public education system with the use of the IQ-achievement discrepancy model to identify children with LD (Fletcher, Coulter, Reschly, & Vaughn, 2004). According to the IQ-achievement discrepancy model, a child is considered learning disabled and provided with learning interventions if standardized testing reveals a significant discrepancy between intelligence and achievement skills (O'Malley et al., 2002). When a learning disability is identified, the child qualifies to receive specialized instruction to increase achievement. If testing reveals that a child has a low IQ (between 70 and 85) with commensurate achievement, the child is labeled a "slow learner," and services to increase achievement are denied. These services are denied based on the intelligence theory which asserts that children who fit the slow learner profile are performing up to their learning potential, and so providing specialized instruction to these children would not benefit them (D'Angiulli & Siegel, 2003; Hall et al., 2005; O'Malley et al., 2002). The current literature does not provide sufficient evidence to support or challenge the intelligence theory and the effectiveness of the IQ-achievement discrepancy model. Furthermore, there is a lack of research on whether RTI is

effective at increasing reading fluency skills. Consequently, many school systems continue to use this model even though the USDOE has developed an optional alternative model for identifying learning disabilities (USDOE, 2004).

Using a quasi-experimental, pre/posttest, control group, quantitative design, this research study investigates the impact of specialized instruction on reading fluency skills with students in Grade 4 who have been identified as at risk for reading failure according to DIBELS Oral Reading Fluency scores. According to the National Reading Panel (2009), "Fluency is one of several critical factors necessary for reading comprehension, and despite its importance as a component of skilled reading, fluency is often neglected in the classroom" (¶ 4). When schools strictly use the IQ-achievement discrepancy model, many students who struggle with reading fluency do not qualify for special education services and therefore are not provided with the academic support they need to be successful, such as supplemental interventions. This lack of academic support often causes these students to experience failure throughout their academic careers (Cooter & Cooter, 2004, p. 57).

Stonovich (2005) described the use of the IQ-achievement discrepancy model this way: "Imagine that your current HMO covered only the procedures and diseases recognized by the medical profession in 1950" (p. 103). Continuing to use this model will allow for more children to fall through the cracks in our educational system (D'Angiulli & Siegel, 2003; Fletcher, Coulter, Reschly, & Vaughn, 2004; Hettleman, 2003; Joshi, 2003; Stage, Abbott, Jenkins, & Berninger, 2003; Stanovich, 2005; Ukrainetz, 2006). A wealth of literature provides evidence suggesting the educational system is in dire need of a mandatory change in the identification of children with learning disabilities, a change that would incorporate learning methods to assist all students who are at risk for failure, including slow learners. Many researchers have argued that the IQ-achievement discrepancy model is ineffective and proposed that all schools systems adopt

a new identification method that will enable slow learners to be successful (Fletcher, Francis, Morris, & Lyon, 2005; Fuchs, Mock, Morgan, & Young, 2003; Stage et al., 2003; Stanovich, 2005). This literature will be discussed in depth in section 2. In order to create this change, more research needs to be conducted to investigate the use of RTI. Furthermore, additional literature in this area will be necessary to support alternative methods of LD identification (Fletcher, Coulter, et al., 2004).

Theoretical Framework

Part of the theoretical framework for this study is the intelligence theory, which asserts that intelligence, or IQ, predicts an individual's ability to learn. This theory, developed by Spearman (1904), is also known as the theory of general intelligence or Spearman's g and is based on the assumption that intelligence is an unchanging characteristic that determines learning capacity (Lubinski, 2004; O'Malley et al., 2002). The theory of general intelligence has been the basis for how the special education system has operated for more than 30 years (Fletcher, Francis, Morris, & Lyon, 2005; Francis et al., 2005; Fuchs et al., 2003). Because the IQ-achievement discrepancy model is based on this theory, it served as part of the theoretical framework for this study

Additionally, the theory of dual discrepancies presented by Fuchs, Fuchs, and Speece (2002) also constituted the theoretical framework for this study. This theory states that "when a low-performing child fails to manifest growth in a situation where others are thriving, alternative instructional methods must be tested to address the apparent mismatch between the student's learning requirements and those represented in the conventional instructional program" (p. 35). This theory, essentially, is the foundation for the RTI model, which was implemented in this

study. Both the intelligence theory and the dual discrepancies theory played vital roles in this study and will be referred to throughout this document.

Purpose of Study

The purpose of this study was to investigate the impact of specialized instruction on reading fluency skills with students in Grade 4 who were identified as at risk for reading failure. These students have been struggling to achieve academic success throughout their educational careers and some have not been given the opportunity to learn through nontraditional means. Many reading interventions have been used with students with LD to increase achievement; however, students who have been identified as slow learners or non-learning disabled based on this model have not been able to participate in such interventions (Cooter & Cooter, 2004; Kavale, 2005; Kavale, Holdnack, & Mostert, 2005). This study attempted to determine if a group of students who are at risk for reading failure was able to increase reading achievement using an evidence-based reading intervention. Because reading rate and accuracy are so closely related to reading fluency, these areas have also been explored in this study.

Significance

This study could be significant to administrators in special education departments who determine qualifications for specialized instruction. This study could also be significant for educators who have students who are at risk for reading failure, to assist in determining methods of instruction and development of lesson plans and curriculum. As stated earlier, according to the National Reading Panel, "Fluency is one of several critical factors necessary for reading comprehension, and despite its importance as a component of skilled reading, fluency is often neglected in the classroom" (National Reading Panel, 2009). Moreover, the USDOE could use the results of this study to aid in the determination of the most effective method of identifying

children with LD, and ultimately implement new special education standards for all school systems.

Aside from the benefits for administrators, educators, and policy makers, this study could also assist in creating social change. Making RTI mandatory in the United States may enable hundreds of thousands of children throughout the United States to receive the reading interventions necessary to experience success in reading. According to the literacy report released by the National Center for Education Statistics (NCES, 2006), about one in 20 adults in the U.S. are not literate in English. With academic support for all students who experience reading failure or risk for failure, the number of illiterate adults in this country will drastically decrease. With this research, along with related research, our society will be equipped with the understanding and tools necessary to guide our educational system so that every child may receive an education that fits their specific needs.

Nature of the Study

This quasi-experimental quantitative study, using a pretest and posttest control group design, investigated the impact of specialized instruction on reading fluency skills with students in Grade 4 who were identified as at risk for reading failure in a large urban community in the northeast region of the United States. The sample consisted of 20 students who were identified as at risk for reading failure through the most recent DIBELS Oral Reading Fluency scores. The instrument used in the study was the Gray Oral Reading Test: Fourth Edition (GORT-4). The materials used were the standard grade level reading curriculum as well as the reading intervention program Read Naturally. The independent variable was the intervention, which consisted of a supplemental individual reading intervention using Read Naturally, and the dependent variable was the students' average GORT-4 scores. The following data were obtained

from school administrators and personnel: students in Grade 4 who have been identified as at risk for reading failure based on the most current DIBELS Oral Reading Fluency scores. Students were randomly dispersed into two groups: (a) those who received supplemental reading intervention using Read Naturally in addition to reading instruction in the general education classroom (b) those who received reading instruction strictly in the general education classroom using the regular reading curriculum. Pre and posttest scores were analyzed using an independent measures *t* test. The researcher did not develop inferences to other populations that are not represented in this study. Additionally, the researcher could not control for the natural maturation, personal experiences, or previous exposure to reading interventions within the sample being studied that may have affected the internal validity of the study.

Research Questions

1. Is there a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group?
2. Is there a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group?
3. Is there a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group?

Hypotheses

*Ho*1: There is no significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*1: There is a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ho*2: There is no significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*2: There is a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ho*3: There is no significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*3: There is a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

The research question and hypotheses will be discussed in depth in section 3.

Definition of Terms

Academic achievement: "the achievement by individuals of objectives related to various types of knowledge and skills" (International Observatory on Academic Achievement, 2006, ¶ 2).

At risk for reading failure: fourth grade students with less than 71 words correct per minute on DIBELS Oral Reading Fluency testing (DIBELS Data System, 2009).

Intelligence: "The ability to learn or understand; the skilled use of reason" (Merriam-Webster, 2009, ¶ 1).

Reading fluency: "The immediate result of word recognition proficiency" (National Reading Panel, 2000, p. 3).

Response to intervention: an alternative, multi-tiered model used to identify students with learning disabilities that "delineates a continuum of programs and services for students with academic difficulties" (National Joint Committee on Learning Disabilities, 2005, p. 251).

Severe discrepancy: at least a 20 point difference between standardized intelligence testing scores and standardized achievement testing scores (LDinfo Publishing, 2007).

Slow learner: a student who has the ability to learn necessary academic skills, but at a rate and depth below average same age peers" (School Psychologist Files, 2008, ¶ 1).

Specialized instruction: training in a specific academic domain (i.e. reading fluency) provided in a small group or one on one setting outside of the general education classroom that is specific to individual student's needs (U.S. Department of Education, 2009)

Delimitations, Limitations, and Assumptions

Delimitations include the following: This study was confined to a relatively small sample of elementary school students in Grade 4 in a large urban city in the northeastern region of the United States. The relatively small sample size was also a delimitation. Participants were chosen by the researcher from a small number of fourth-grade students. One limitation of this study was that DIBELS Oral Reading Fluency scores may have been obtained at various times. Other limitations included the type of specialized instruction provided, which varied depending upon students' needs. Because of the small sample size, the findings may not be generalized to any other population. Furthermore, this study assumed that the researcher was not biased towards any of the participants based on their educational history or any other characteristics or qualities.

Summary

As school systems across the United States remain in flux in the identification of specific learning disabilities, many children continue to suffer the consequences. The data obtained in this study contributes to the literature, and could ultimately assist school systems in making the decision to use the most effective LD identification method in order to allow all students to receive an education that meets their individual learning needs. The field of education must

continuously undergo transformations in order to provide students with the knowledge they need to become productive members of our world. One way to achieve this evolutionary goal is to continue conducting research on the cutting edge issues in education. The identification methods debate has been ongoing over the past two decades, and continues to be unresolved. This and future research could help educational leaders to promote social change by coming to a consensus on how to best solve this longstanding problem so that we may move forward and continue to make advancements in the field of education, ultimately advancing our community.

Section 2: Literature Review

Introduction

Effective educators possess a deep-seated passion to enrich the lives of children by guiding, motivating, inspiring, and shaping their future. This task requires patience and empathy along with strong intuition and an understanding that every child is unique. No two children have the same strengths, weaknesses, and learning style. In order to foster success for every student, educators must modify their teaching methods and approaches to meet the learning needs of each student through experimenting with various interventions (Sternberg, 2006). Ideally, this is how the educational system should operate. Unfortunately, this is not the reality (Sternberg, 2006). In many cases, students' needs exceed what a general education classroom teacher is able to provide, and neither the student nor the teacher are given supports to address their needs (Cooter & Cooter, 2004).

This study was conducted to provide children who experience difficulty learning through traditional means an opportunity to succeed. Historically, many of these students have not been given a chance to increase their achievement skills, even though they struggle with the curriculum, because they do not qualify to receive scientifically based interventions through special education (Ukrainetz, 2006). As stated in section one, even though the USDOE has created new guidelines for identifying learning disabilities, more specifically the RTI method, many school systems are maintaining the use of the IQ-achievement discrepancy model, and consequently children continue to be denied services to increase achievement.

This literature review analyzed and synthesized the current research on IQ measures, the IQ-discrepancy model, response to intervention, and reading achievement. Section 2 also presents information on the historical background of intelligence and the intelligence theory as

well as RTI, which provide the theoretical framework for this study. All relevant literature was obtained using research databases, including academic search premier, education research complete, educational resource information center (ERIC), health and medical complete, health sciences, journal of the American Medical Association, medline, psychARTICLES, psychBOOKS, psychology, and psychINFO. A variety of key words were used to search these databases, including learning disability, slow learner, response to intervention, low IQ, intelligence, reading intervention, and reading fluency.

Historical Background

Nettelbeck and Wilson (2005) discussed the history of IQ testing, changes that IQ theories and testing have undergone, the impact of IQ testing in education, and implications for the future. The authors asserted that IQ has been a controversial issue for centuries, and the debate continues to grow as researchers discover more about IQ and intelligence.

The first IQ test was developed by Alfred Binet in the early 20th century to identify school children who could be expected to face challenges learning the curriculum. Since that time, IQ testing has undergone many changes. Following Binet, Spearman (1927) developed a theory based on his research that IQ can be narrowed down to a single generalized ability measure, which he labeled Spearman's g. Spearman's g has been widely accepted, according to Nettelbeck and Wilson, throughout Western European and North American cultures. More recently, however, Gardner (1983, 1999) has been on the forefront of IQ development with his theory that "human cognitive abilities are best envisaged as several independent forms of intelligence" (p. 615). The authors indicated that such theories of multiple intelligences date back to Thorndike (1926) and Thurstone (1938), who played critical roles in educational psychology. Pioneers in multiple intelligence theories have been vital in education because the "assessment

system that can identify relative strengths and weaknesses in a profile describing a child's cognitive performance has the potential to generate practical interventions, whereas a global IQ score offers little opportunities for this" (Nettelbeck & Wilson, 2005, p. 615). In the realm of education, this type of IQ analysis is more practical and useful, and provides educators with guidance in terms of appropriate interventions.

In 1993, Carroll broke intelligence down even further by identifying specific domains that form IQ; however, according to Nettelbeck and Wilson, "commonality among these [domains] in turn defined a single general factor, which Carroll equated with Spearman's g" (p. 616). Following Carroll's work in the field, Gardner went on to describe different activities that comprise intelligence, aside from the traditional areas, such as music, dance, and kinesthetics. There is still much debate over Gardner's theories, and Nettelbeck and Wilson pointed out that Gardner's theories are strictly a matter of opinion.

Sternberg has also moved beyond the traditional views of intelligence and collaborated with other researchers over the past 30 years to develop several intelligence theories as well as tests that measure different constructs of intelligence. Sternberg "acknowledges that abilities including creativity will rely to some degree on inherited characteristics, and Nettlebeck and Wilson agree" (Nettelbeck & Wilson, 2005, p. 619). Sternberg continues to develop his theories through research and remains pioneer in education to this day. His theory of mental self government is be discussed later in this section.

Nettelbeck and Wilson went on to discuss the future of IQ testing and intelligence. The authors posited that, as with all things, theories will continue to develop and change over time. Furthermore, the methods of IQ testing will most likely change as well. The foreseeable future of cognitive testing will be directed towards applying theory to the improvement of existing

instruments, as well as developing new tests. A primary characteristic of future tests will be that they will assess a wide range of cognitive domains, as well as general ability (p. 623).

The authors also believe that the uses of IQ testing will expand over time; however, in the realm of education, IQ testing will always be a single tool in the assessment of students, and other forms of assessment should be used in such situations.

As mentioned earlier, in 1997 Sternberg created an intelligence theory known as mental self government (Sternberg & Zhang, 2005). Since that time, Sternberg as well as his colleagues and other scholars have expanded upon and conducted studies to support this theory (Fer 2005; Park, Park, & Choe 2005; Sternberg & Zhang 2005; Zhang 2002). Sternberg based his theory on the idea that individuals operate in a similar structure as the branches of government, which he divided into several key components including functions, forms, levels, scopes, and leanings. Each component is broken down even further to incorporate a variety of thinking styles, which are "preferred ways of expressing or using abilities" (Fer, 2005, p. 56).

According to Sternberg, the functions of government consist of legislative, executive, and judicial. The legislative student prefers creating, planning, formulating, and strategizing independently. The executive student operates best with strict guidelines of structure, procedure, and rules to abide by. Finally, the judicial student learns best with tasks that require using analytical skills. According to Sternberg and Zhang (2005), the executive student performs best with multiple choice and short answer tests as well as with tasks that require rote memorization.

Sternberg broke the forms of government down into four distinct types including monarchic, oligarchic, hierarchic, and anarchic. The monarchic student learns best when focusing on one task at a time until completion, whereas oligarchic students prefer working on multiple tasks within a given period. Hierarchic students tend to prioritize tasks and set goals for

themselves to complete in a specific order. The anarchic pupil has difficulty adhering to structure and rules. This type of student prefers to operate in a more flexible manner.

Sternberg went on to separate levels of government into two distinct categories, local and global. Local learners prefer tasks that utilize specific and concrete details whereas global learners perform best on tasks that require extrapolating abstract ideas. According to Sternberg and Zhang (2005), "Local students enjoy tasks that require them to keep track of details and to focus on concrete specifics of a situation" (p. 248). Much like the levels of government, the scope of government is also comprised of two distinct types, internal and external. According to this theory, students who are internal learners prefer working independently and are uncomfortable interacting with peers in a group, whereas external students prefer group work and projects. Sternberg categorized the final aspect of the self-government model, leanings, into liberal or conservative. Liberal students perform well on tasks that are unfamiliar, ambiguous, and challenging. Conservative pupils enjoy familiar clear-cut tasks, and they avoid change at all costs.

According to Sternberg's theory, the probability of a student achieving success will increase if teachers accommodate the child's thinking style, which includes varying the types of assignments, projects, and evaluations the child takes part in. Kim (2005) held similar views. After conducting an extensive literature review, spanning from 1961 to 2004, on the subject, Kim discovered that the "negligible relationship between creativity and IQ scores indicates that even students with low IQ scores can be creative. Therefore teachers should be aware of characteristics of creative students – this will enable teachers to see the potential of each child" (p. 65). Sternberg's intelligence theory, as well as Kim's beliefs, contradicts Spearman's g in that he believes intelligence is not a single entity, but that it takes on various forms. Sternberg has

asserted that intelligence cannot be measured in a narrow assessment, rather it needs to be examined on a more creative and practical level. Spearman's g constitutes a single construct and can be measured as such. Both of these theories continue to be used in education today; however, both are fundamentally different.

Gardner and Moran (2006) expressed similar beliefs about multiple intelligence theory. They asserted that "the debate of whether intelligence is a singular individual quality [g] or a plethora of components has waxed and waned throughout the 20th century" (p. 227). While Spearman believed that intelligence is in fact a single measurable entity, Gardner believed that a variety of intelligences exist. Through his research, Gardner has discovered eight separate and distinct intelligences: linguistic, logical-mathematical, musical, spatial, bodily kinesthetic, naturalistic, interpersonal, and intrapersonal. From a historical perspective, Gardner and Moran asserted that,

> It has taken a century and many millions of dollars to bring IQ testing to its current, not especially impressive status. IQ-based education has yielded no genuine implications for classroom instruction or the improvement of performances, only feelings of pride or shame on the part of test-takers and their families. (p. 229)

According to these researchers, the multiple intelligence theory would much better suit students in the educational setting than traditional intelligence measures. Using this theory, teachers would be able to capitalize on students' strengths in order to educate them effectively.

Although Gardner and Moran make a compelling argument for this theory, there is still no way to test the various intelligences. The lack of formal assessments to determine students' intelligences would make it difficult for the educational system to subscribe to. Furthermore, the debate about intelligence as a single entity or various components is ongoing, and more research needs to be conducted to support the multiple intelligence theory.

In 2007, Rindermann examined intelligence on an international level when he developed an article on international cognitive ability, in which he reviewed and analyzed actual IQ data obtained from 113 different countries to determine if student achievement and assessments largely measure a common cognitive ability. The data were originally collected by Lynn and Vanhanen (2002, 2006). Rindermann identified the major differences between cultures that may cause variations in data, such as attendance rates, differences in grade levels in terms of student ages, sample representativeness, and variability among definitions of intelligence. Rindermann stated that, "International student assessments and intelligence test studies suffer from different problems of sample representativeness and statistical methods; mainly the comparability across nations is questioned" (p. 672). Due to this instability, adjustment formulas had to be used to gain a clearer and more accurate understanding of IQ across countries, which may have compromised the accuracy of the data overall. Furthermore, of the 113 countries, IQ data had to be estimated for 79 due to missing information. Once all of the data were adjusted, the researcher examined the information closely.

Rindermann analyzed data from several different sources, including International Assessment of Educational Progress (IAEP), Programme for International Student Assessment (PISA), Progress in International Reading Literacy Study (PIRLS), Trends in International Mathematics and Science Study (TIMSS), and various intelligence tests. After reviewing the testing statistics, Rindermann discovered that correlations among all of the above mentioned assessments were high, with the lowest being $r = .91$ in the IEA reading study. Moreover, a factor analysis for all data revealed "g-factor explained 94% (unadjusted) or 95% (adjusted) of the variance of the student assessment scales and the intelligence tests" (p. 681), thus "Different

scales of student assessment studies and different cognitive test approaches appear to have measured essentially the same construct, namely general national cognitive ability" (p. 687).

The results of this study, as well as the results from Lynn and Vanhanen's previous studies provide support for spearman's g. Rindermann noted that, "Intelligence is a term frequently used in occidental history of thought. It was used in the Middle Ages…., in the Age of Enlightenment, and since the 19th century, it has been the well-established term for the *ability to think*" (p. 692). According to this study, throughout history and across nations, the intelligence construct, has been and continues to be centered around a single generalized ability measure. This is somewhat contradictory to Nettelbeck and Wilson's (2005) article, where they asserted that the intelligence paradigm has evolved throughout the years and controversy continues to surround this subject.

Flynn and Weiss (2007) also provided a historical perspective of IQ with regard to how IQ levels have changed between 1932 and 2002. They reported findings from previous studies conducted by Flynn (1984) which analyzed IQ data from 73 studies and over 7,500 subjects from 1932 to 1972 as well as 26 studies covering 2,226 subjects from 1972 to 1989. In addition, Flynn and Weiss used WISC-III and WISC-IV data obtained from 244 individuals provided by the Psychological Corporation. Results revealed that between 1932 and 2002, Americans have demonstrated a total gain of almost 22 IQ points. According to the authors, the data they analyzed show that "today's schoolchildren take on-the-spot problem solving more seriously and are more comfortable with the vocabulary of science" (p. 222). Like Nettelbeck and Wilson's (2005) study, Flynn and Weiss' study indicates that IQ is evolving over time.

Although Flynn and Weiss presented some compelling data, the statistics evoke some skepticism. IQ tests, educational demands, and societal norms have changed over time, which

some may argue could have constituted the changes in IQ over time. These transformations make it difficult to make comparisons among students' performance in 1932 versus how students performed in 2002. Consequently, one might ask, is the g-factor truly evolving, or is this change due to extrinsic factors, and the g-factor is really remaining stable over time? This is a difficult question to answer without additional research.

White discussed the historical background of intelligence and the g-factor in his 2005 article, "Puritan Intelligence: the Ideological background to IQ." White described the works of many philosophers and psychologists and discussed how their theories have evolved over the years. He began by introducing Galton, cousin of Charles Darwin, who first published his beliefs about intelligence in 1865. Galton labeled his idea "eugenics," which entailed "encouraging our finest intellectual stock to breed and discouraging our meanest from doing so. The idea that there are innate, and limiting, differences in intellectual ability – is the basic assumption in this proposal" (p. 424).

Eris (2008) also discussed the work of Galton. He indicated that during that time period, IQ testing was used as a, "simple economical but strong mechanism which classified individuals on the basis of their levels of intelligence, determined their entry into the country on that basis, and specified the kinds of jobs they were capable of doing" (p. 81). Eris went on to state that "By testing the immigrants, the state ensured control in a unidirectional way…and gave the message to those who failed the test that they were inferior and powerless" (p. 81).

Following Galton, many scholars began conducting their own work in the area of intelligence and advocating their use for various reason. Burt believed that intelligence testing should be used to help decide the educational placement of children, where Goddard (1913), the first intelligence tester (White, 2005, p. 433), felt that intelligence testing should be used to test

immigrants coming to the United States through Ellis Island to "exclude those of poor intellectual stock" (p. 424). Terman (1919), who created the Stanford-Binet test of intelligence (p. 435), proposed that intelligence tests would be best used to make vocational decisions. McDougal and Pearson (1930) attempted to devise a set of intelligence tests that were independent of language and learning that would be universally applicable, and believed that "social progress would inevitably favour [sic] those who worked mainly with their brains rather than their hands" (p. 433).

White went on to discuss Spearman's g, or general intelligence, and Thomas, who felt that intelligence dealt with abstractions, concepts, relationships, and logic. White claimed that "to fully understand this, we have to grasp the crucial importance of logic in the Puritan world view," where he posited the roots of intelligence lie (p. 437). White concluded that all of the above mentioned scholars demonstrated Puritan connections in their intelligence theories.

Nash (2005) disagreed with the concept of Spearman's g, stating that

> The discussion of human capabilities requires three distinct concepts. First, there is the concept of potential to develop and learn…Secondly, there is the concept of skill as a state of being that enables one to perform…And thirdly, there is the concept of ability as the capacity to carry out certain skilled actions. (p. 8)

Rather than intelligence being a single measurable entity, Nash asserted that intelligence is comprised of these three concepts. Furthermore, Nash believed that IQ was irrelevant to educational success, stating "There is, indeed, no way of knowing what an individual with a given IQ score can achieve in the educational system" (p. 10).

Greenspan (2006) concurred with Nash's assertions in his discussion of the label mental retardation. Greenspan indicated that the terms "educable" and "trainable" are still used today to describe "what someone can do functionally" (p. 210). The author went on to state that the main distinction is, "the ability to benefit from reading instruction (the problem, however, is that these

have been hijacked by IQ criteria, with the result that many children who could learn to read have been wrongly shunted into trainable classes)" (p. 210). These terms that continue to be used in education, as Greenspan pointed out, have roots in Galton's theory.

Whether intelligence is a single measurable phenomenon or a multifaceted phenomenon has yet to be determined. It is conceivable that the debate will continue indefinitely. Or perhaps future research will uncover the truth. Theoretical connections in intelligence and the response to intervention model, as well as the implication of these theories on reading disabilities, will be discussed in the following sections.

Theoretical Connections

Intelligence

Glutting, Watkins, Konold, and McDermott (2006) conducted a study to explore the relative importance of specific versus general abilities from the WISC-IV (Weschler Intelligence Scale for Children- fourth edition) in predicting reading and mathematics achievement. According to the researchers, "subtest analysis has undergone serious challenges over the past 2 decades, both methodologically and empirically," (p. 103) and the authors contributed to the research by examining this issue.

Glutting et al. collected data from 498 participants ranging in age from 6 years 0 months to 16 years 11 months. The researchers obtained standard scores from both the WISC-IV and the WIAT-II (Weschler Individual Achievement Test), and analyzed the scores to determine how the observed intelligence scores contributed to the prediction of achievement scores. The researchers discovered that only the g and verbal comprehension factors influenced reading and mathematics scores in terms of model fit and parsimony. These results impact how psychologists interpret test scores in that "some psychologists have recommended interpretation of factor indexes over the

FSIQ [full scale IQ] when predicting achievement; however, to the contrary, current results reveal that psychologists need to interpret only the FSIQ to predict performance in reading and math" (p. 111). Furthermore, "professionals might have many legitimate reservations about using IQ-achievement discrepancies to diagnose LD" (p. 109). Typically, psychologists look beyond the Full Scale IQ score to determine a student's strengths and weaknesses, and to make recommendations. Interpreting only the FSIQ would not allow psychologists to analyze the scores or make the necessary recommendations. Although, when attempting to identify a learning disabilities, psychologist and special education teachers compare the FSIQ with scores on achievement testing to determine if a discrepancy exists.

D'Angiulli and Siegel investigated the importance of IQ testing in identifying learning disabilities in their 2003 study. The authors believed that "IQ tests are not particularly useful for determining who has a learning disability, and findings from other studies suggest that IQ test scores are not particularly useful for determining who will benefit the most from remediation" (p. 56). D'Angiulli and Siegel conducted a study involving 364 students, ages 7-16, with IQs of 80 or above who were classified as either reading disabled, arithmetic disabled, or typically achieving. The examiners administered the Wechsler Intelligence Scale for Children-Revised (WISC-R) as well as various achievement tests, and analyzed the performance patterns for each participant. Results indicated that the patterns on intelligence tests were not a reliable measure of learning disabilities, and patterns on the achievement tests were a reliable measure of learning disabilities. Based on their study, D'Angiulli and Siegel claimed that, "achievement scores are still the most satisfactory measures for identifying reading and learning disabilities regardless of IQ level" (p. 54).

Results from D'Angiulli and Siegel's study should be interpreted with caution because the information gathered pertains to only one intelligence measure. Psychologists do not typically use one standardized test for every student they assess, but rather they have access to a variety of measures to choose from based on students' ages, grade level, and general abilities. As Fletcher et al. (2004) pointed out, "any test has measurement error" (p. 316). Moreover, "no single score can perfectly capture a student's ability in a single domain" (Francis et al., 2005, p. 100). Conducting a similar study using additional intelligence measures would make D'Angiulli and Siegel's claims more reliable.

Another factor that should be kept in mind when interpreting the results is that D'Angiulli and Siegel's subjects consisted of children with IQ scores higher than 80. The subjects' IQ scores may or may not have accounted for or influenced their patterns of performance. In addition, the outcome of this study, which suggested that achievement scores are the best indicator of a learning disability, does not have implications as to how a student will respond to interventions addressing the deficit areas.

Fuchs and Young (2006) contributed to the literature by examining the current data on the use of intelligence tests, which resulted in mixed conclusions. Some of their findings are contradictory to Al Otaiba and Fuchs' results, while others were not.

In their review, the authors utilized 13 studies consisting of 1, 542 subjects who were at risk or reading disabled. Subjects in 6 out of the 13 studies participated in phonological awareness intervention, and subjects in the remaining 7 out of 13 studies participated in more comprehensive intervention consisting of decoding, fluency-building, or reading comprehension, in conjunction with phonological awareness training. Analysis of the results indicated that IQ was relevant in 3 out of 6 phonological awareness interventions, and in 5 out of 7 studies on

comprehensive reading interventions. The overall impression was that IQ was relevant in 8 out of 13 total studies.

Fuchs and Young examined studies based on strict criteria, including the following:

1. Studies were taken from peer-reviewed journals.

2. Examiners administered intelligence tests or proxies for an intelligence test.

3. Examiners provided research-based intensive reading interventions.

4. Examiners used reading or related measures to identify the effects of the interventions.

5. Subjects included children that were at risk for reading failure or reading disabled.

Fuchs and Young's study contained variables that may have impacted their conclusion. The studies that Fuchs and Young analyzed were variable in structure, methodology, and subject-base. So much variability among each study may have caused the information obtained to be less comprehensive.

In addition, Fuchs and Young analyzed studies that were conducted at various points between 1993 and 2003, which is a significant time span. By 2006, when their article was published, the 1993 studies were not up-to-date. Furthermore, the IQ measures and the type and intensity of interventions were extremely variable. In some studies, the examiners administered the McCarthy Scales of Children's Abilities, and in other studies, the examiners administered The Wechsler Intelligence Score for Children-Revised, which may have resulted in different outcomes. The most significant discrepancy among the studies, however, was the variability among interventions. The total duration of the interventions that were implemented ranged from as little as 4 hours to as much as 68 hours. Making comparisons of the subjects' performance across such variable conditions is difficult and unreliable. Researchers should compare results from studies with similar conditions in order to make reliable conclusions.

Edwards (2006) developed an article addressing IQ and learning disabilities form a multicultural perspective. The author indicated that students of various ethnic groups are disproportionately placed in special education due to the use of intelligence tests that do not take mean IQ differences among ethnic groups into account when being normed. Edwards indicated that when unadjusted for differences in socio-economic status, on average "individuals of Asian descent scored higher than those of European descent, individuals of African descent scored lower than those of European descent, and individuals of Hispanic descent scored somewhere in between" (p. 246).

Because tests are not normed by groups, challenges arise when examiners choose which test to use for students from various ethnic backgrounds. Edwards pointed out that the Standards for Educational and Psychological Testing specified that this information should be taken into consideration when determining which test to use, which places psychologists in a difficult situation. Furthermore, because this information is unavailable and therefore not considered, many students are being incorrectly placed in special education programs. Edwards concluded by suggesting that research continue to be conducted on "the goodness-of-fit between test instruments and what they are supposed to measure and for whom" (p. 248).

Following his 2006 article, Edwards conducted a study in 2007 expanding upon this issue. Edwards reviewed previously published reports to address the factors associated with the overrepresentation of various ethnic groups in special education. Edwards pointed out that "scores from intelligence tests remain an integral part of MR [mental retardation] and DD [developmental delay] educational diagnoses (IDEIA, 2004) and often carry significant implications for special education and placement" (p. 31). Therefore, the author stressed the importance of obtaining fair and valid measures when administering IQ testing. Edwards

gathered information regarding the representation of various ethnic groups receiving special education services, and found that African American and Native American students were represented at a higher rate than all other ethnic groups. The author reported that "African American students have a 200% higher probability of placement in the MR category and a 65% higher probability of placement in the DD category than other groups combined" (p. 33).

Due to this staggering statistic, the author conducted research to determine potential causes for this overrepresentation and discovered that philosophical orientation, testing norms cycles, and differences in means among ethnic groups contribute to the disproportionate numbers. Edwards also pointed out that intelligence test development utilizes varying intellectual or psychological theories, and tests are created with emphasis on different areas (i.e. some are language loaded while others are culturally dependent.).

Edwards concluded that this may cause "students to be passed over for services (false negatives) and students to be inappropriately identified as needing services (false positives), consequently being unfairly and inequitably treated by the educational system (Donovan & Cross, 2002)" (p. 34). Edwards indicated that this may lead to a significantly flawed testing process when identifying students with special needs.

Contrary to Edwards' study focusing on overrepresentation of ethnic groups in special needs programs, Fletcher-Janzen and Ortiz (2006) developed an article to discuss the role of IQ testing in identifying gifted and talented, culturally and linguistically diverse children, and to make recommendations for fair and equitable assessments of these children. The authors asserted that children from diverse cultures consistently score lower on tests of intelligence than do Anglo/Euro Americans due to several factors, including (a) language barriers (b) cultural

differences (c) bias in the development of testing items (d) socioeconomic status (SES), and (e) cultural loading in IQ tests.

The authors reviewed a study by Fletcher-Janzen and Daniel (2005) in which 6 culturally diverse groups of children were administered two different IQ tests (Kaufman Assessment Battery for Children – Second Edition and Weschler Intelligence Scale for Children – Fourth Edition). Four of the 6 groups were from middle to high SES backgrounds and the remaining 2 groups were from low SES backgrounds. Fletcher-Janzen and Daniel found that the middle to high SES group scored average to above average on both tests, and the low SES group scored significantly lower on the WISC-IV than the KABC-II. These data are alarming as the WISC-IV is widely used among psychologists to identify learning disabilities. Fletcher-Janzen and Ortiz suggested that

> it may well be that the results indicate that the WISC-IV, which represents a traditional hierarchical view of intelligence with global ability or "g" at the apex, is an appropriate measure of intelligence for ethnic minority children but only for those who come from middle to high SES backgrounds (p. 129).

These findings suggest that intelligence testing may not accurately identify the learning potential of students from diverse backgrounds. Fletcher-Janzen and Ortiz have indicated that such students frequently score lower than their true ability. Even though the authors are addressing this issue from the standpoint of identifying gifted and talented students, this also has severe repercussions for students who are struggling with the curriculum and are being tested for potential learning disabilities. Given this information, it is likely that these students' IQ scores may be commensurate with their low achievement scores, thus these students would be denied

special education services due to lack of evidence of a learning disability using the IQ-discrepancy model.

Fletcher-Janzen and Ortiz have suggested that practitioners develop a new set of standards when identifying learning potential and talents that include (a) developing self-awareness (b) establishing rapport and building trust (c) learning the family system (d) rethinking intelligence (e) modifying referral indicators (f) evaluating fairly (g) interpreting equitably, and (h) intervening appropriately. The authors indicated that tapping into different areas of intelligence testing this way, rather than through traditional means, will give practitioners a more comprehensive view of learning potential and ability.

Epelbaum (2007) conducted a case study investigating the use of multiple intelligence theory on reading comprehension with an English language learner (ELL). The author claimed that "Comprehension is one of the most difficult skills to assess well and within the actual classroom; it is much too easy to overlook individual struggling students" (p. 243). Epelbaum believed that if a teacher taps into a student's strongest area of intelligence, the student will be better able to learn the given material and improve their level of skill in that area. The subject of this case study, a tenth grade student from Bangladesh, was chosen to participate in this study due to her weak reading comprehension skills. Epelbaum began by identifying the subject's strongest intelligences through observation, interviews with teachers, and the completion of a multiple intelligence survey and checklist. She then created comprehension assessments addressing the subject's two strongest intelligences, which were interpersonal intelligence and verbal linguistic intelligence. The subject then participated in a variety of reading comprehension assessments, including the assessment that was created based on her intellectual strengths. The researcher discovered that not only did she perform better on the assessments that addressed her

strongest intelligences, she also enjoyed those assessments the most. During these assessments, Epelbaum noted that "She was more engaged, less tired, and she even smiled and laughed here and there" (p. 249).

Epelbaum's study supports Gardner's argument that individuals learn best when teachers understand and cater to their intellectual strengths and weaknesses. However, this was a single subject study that cannot be generalized to other populations. Furthermore, the subject's lack of exposure to the English language may have impacted the results of this study. More research on multiple intelligences needs to be conducted using a larger population of students as subjects.

Just as educators and psychologists use IQ testing to identify learning disabilities, IQ testing is also used to identify language disorders. Dethorne and Watkins (2006) were interested in determining to what extent language abilities are associated with nonverbal IQ among children with language disorders. In order to make that determination, Dethorne and Watkins conducted a study in which they administered several standardized IQ tests, standardized language assessments, and informal evaluations to analyze the correlation between the two domains. The authors asserted that "a child whose standardized language scores fall two standard deviations below the mean and whose IQ score is similarly low might not be eligible for language intervention, and might instead be designated as a 'slow learner'" (p. 641). They pointed out that some empirical research has shown "language skills can be improved despite lower IQ" (p. 643).

Participants consisted of 30 children with language impairment between the ages of 4 and 8. Each child underwent 2 to 3.5 hours of testing and assessments on an individual basis. The researchers then compared the results and found that non verbal IQ accounted for 20-21%, at most, of the variance in language ability. Although this small correlation exists, according to the authors, it cannot be assumed that language development is due to non verbal IQ. According to

Dethorne and Watkins, previous studies on treatment efficacy have demonstrated that children are able to respond to intervention regardless of the presence of a language-IQ discrepancy; "Lower IQ cannot be assumed to limit language learning capabilities" (p. 654). The authors suggested that further research needs to be conducted in this area to support this paradigm.

This study is eye opening in that it expands the population of interest beyond students with potential learning disabilities. The use of IQ testing also impacts students who potentially have a language disorder as well. This Broadens the population of students effected by IQ scores, and reveals that IQ testing is more prominent in education than many may realize.

Hoskyn and Swanson focused on the identification of reading disabilities in their 2000 study. The researchers conducted a study in which they synthesized the published literature comparing students with IQ-reading discrepancies to students who are poor readers with commensurate IQ and reading scores. They hypothesized that children in both categories will have similar cognitive deficiencies specific to phonological processing. The authors asserted that "The implicit assumption for the inclusion of discrepancy scores in the classification of LD is that children who experience reading, writing and/or mathematics difficulties, unaccompanied by low IQ, are distinct in cognitive processing from low achievers…which is equivocal" (p. 102).

Hoskyn and Swanson located a total of 69 studies and used specific criteria to determine which studies to use in their synthesis. A total of 19 studies met the criteria. The authors coded the studies in terms of characteristics, demographics, magnitude of discrepancy, mean IQ, and standard deviations on cognitive processing tasks (which were further coded into four categories). The authors then calculated and analyzed the effect size for each dependent measure, and compared one model against another in predicting effect size.

In their analysis, the researchers highlighted several important findings, including (a) phonological processing was comparable between the two groups (b) both groups are also alike in non-phonological measures (c) a common general processing deficit was shared between the two groups (d) syntactical knowledge is a major variable on which the RD children have an advantage over LA children. Hoskyn and Swanson concluded that their synthesis "concurs with several individual studies indicating that the discrepancy between IQ and reading is not an important predictor of cognitive differences between LA children and RD children" (p.115).

The findings from this study were significant for school psychologists, special education teachers, and other personnel who determine eligibility for special education services. However, there are also limitations to this study. The samples primarily included males from upper to middle class settings, the studies that were analyzed included a limited IQ range, and they only analyzed studies that used discrepancy criteria. Because of the limitations in this study, it is difficult to generalize the results. More research needs to be conducted to accounts for such limitations.

Lassiter, Matthews, and Feeback (2007) conducted a study to investigate the validity of a specific intelligence measure known as the CTONI (Comprehensive Test of Nonverbal Intelligence) by comparing it with another well known and widely used cognitive measure known as the WJ-III COG (Woodcock Johnson III Tests of Cognitive Abilities). The authors claimed that very few studies have been conducted to examine the validity of the CTONI, as it is a relatively new instrument. This instrument was designed to measure intelligence without requiring the use of language, to accommodate for individuals from various ethnic backgrounds as well as individuals who exhibit language limitations.

Participants in this study consisted of 60 adults between the ages of 21 and 51 who were recruited from undergraduate and graduate psychology courses at southeastern university. The CTONI as well as selected subtests from the WJ-III COG were administered to each participant in a university laboratory setting by advanced graduate students. The results of each test were then compared, revealing that "the CTONI score might not accurately measure the intellectual functioning of individuals who are intellectually gifted or of low average intelligence" (p. 575). Based on these results, Lassiter, Matthews, and Feeback recommended that psychologists interpret scores obtained from the CTONI with caution.

The results from this study are compelling, in that individuals with certain intellectual characteristics may be misrepresented. However, there were some deficiencies in this study. The CTONI was compared to various subtests from the WJ-III COG rather than the entire battery. In addition, the CTONI was developed to measure a specific type of intelligence known as fluid intelligence, whereas the WJ-III COG was developed to measure both fluid as well as crystallized intelligence (Lassiter, Matthews, & Feeback, 2007). Therefore, this comparison may not have yielded completely accurate results.

Conversely to Lassiter, Matthews, and Feeback, (2007) investigation of nonverbal IQ measures, Runco, Dow, and Smith (2006) examined IQ from a language standpoint. The authors conducted a study based on their belief that IQ measures consisting of divergent thinking tasks

> are bias in the sense that test scores indicate how much experience an examinee has and not only how much thinking ability. This implies that disadvantaged students may be unfairly penalized if they do poorly on tests of divergent thinking as poor performance might reflect their lack of experience rather than low ability (p. 270).

Runco, Dow, and Smith conducted a study to determine the role of experience, knowledge, and memory in divergent thinking tasks using 115 participants who were undergraduate students at the university of southern California.

The participants in this study were administered open-ended tests that allowed original responses as well as fact based tests. After analyzing the participants' performance on the tasks presented, the researchers concluded that some divergent thinking tasks can be significantly influenced by an individual's knowledge. Without specific experiences, certain individuals were unable to provide the required language based responses to answer specific questions.

The results of Runco, Dow, and Smith's study imply that students who lack life experiences are likely to perform more poorly on certain IQ tasks than students who have rich life experiences. This poses a significant problem, particularly for students of low socioeconomic status. However, there were several imperfections in this study. The tests that were administered were not specified by the authors, the incidence of divergent thinking tasks in IQ tests were not discussed, and the questions the participants were asked were limited to one experience.

Although IQ testing has traditionally been used to identify various types of learning disabilities, the U.S. Department of Education has adopted an alternative method of identifying learning disabilities, known as response to intervention (RTI), which does not require the use of IQ testing. Many individuals in the field have argued that it is unethical to use this new method because the definition of a learning disability specifies a significant discrepancy between IQ and achievement (Kavale, 2005). However, others are embracing the new model and feel that regardless of the definition, educators will still be able to identify learning disabilities using RTI (Barnett, Daly, Jones, and Lentz, 2004; Compton, Fuchs, Fuchs, and Bryant, 2006).

Response to Intervention

In light of the reauthorization of IDEA (2004) to incorporate the response to intervention approach as an alternative model of identifying children with learning disabilities, Barnett, Daly, Jones, and Lentz (2004) developed an article presenting implementation guidelines for RTI. The authors described RTI as a process that "builds upon concepts found in IDEA as well as the No Child Left Behind Act….which requires that students undergo effective instruction and progress monitoring before entering special education, to provide a starting place for educational accountability" (p. 66). In their guidelines, they suggested that educators identify children that are not learning the curriculum at the same rate as their peers and provide them with interventions based on their individual needs. They indicated that students should begin with a low intensity intervention (tier one) and move to a more intense level of intervention (tier two) if progress is not made. Barnett, Daly, Jones, and Lentz advocated that special education teams use single-case designs to examine the intervention's effectiveness, and they suggested that students should qualify for special needs if they do not respond to the first two tiers of intervention. The authors provided two hypothetical case studies showing how the students should move through this process.

Compton, Fuchs, Fuchs, and Bryant (2006) also discussed the use of the response to intervention model and took their analysis one step further than Barnett, Daly, Jones, and Lentz by pinpointing possible difficulties in identifying candidates for tier two intervention. They conducted a longitudinal study to determine what individual tests are the most effective in identifying children that are at risk for reading disabilities (RD), and what type of statistical analysis is more accurate in predicting students with RD. Their subjects, 206 children from urban and suburban districts, were followed from the beginning of first grade to the end of second grade. The researchers discovered that word identification fluency, phonemic awareness, rapid

naming skills, oral vocabulary, and short term progress monitoring are the most effective screening tests. Based on their results, the authors recommended that there be "some general agreement about how we define RD," (p. 408) and decision rules must be created to help educators determine what criteria need to be met to qualify a child for tier two intervention.

Gerber (2005) addressed potential obstacles of implementing RTI, as well as discussed the teachers' role in the process. The author pointed out that experienced teachers possess an innate ability to gauge a student's capacity to respond to different levels of instruction. However, he also pointed out that teachers can only expend so much energy on one student before accessing additional resources becomes necessary. Gerber surmised that "when students are significantly different in response to cannot achieve optimal outcomes for each and every student without significant and relevant new resources" (p. 519). The author went on to state that "RTI, therefore, when implemented according to rigorous and standardized protocol, aspires to be a reliable and valid measurement of learning disability" (p. 519).

Even though RTI is promising, Gerber indicated that using the additional resources needed to implement RTI effectively would cause financial strain on school systems. Gerber provided an outline of the cost estimation, suggesting that the national cost would be in excess of 2 billion dollars with teacher student ratios of 1:5 and 1:3, making it difficult, if not impossible, for many school systems to implement effectively. Gerber based his estimated costs on various sources of national data. This, Gerber pointed out, is just one of the many obstacles the educational system will face in implementing this new approach.

Mastropieri and Scruggs (2005) discussed the implications of RTI and the impact on teachers and specialists within the context of RTI. They explored issues of implementing RTI with consistency, RTI as it relates to the No Child Left Behind act, the impact RTI will have on

the definition of learning disability, the role of cognitive measures in assessment, and where the responsibility will lie in ensuring RTI is implemented correctly. They argued that "before these challenges can be met, the field needs to fully operationalize what is meant by the RTI model" (Mastropieri & Scruggs, 2005, p. 526).

As mentioned earlier, the basic structure of the RTI method involves different levels or "tiers" of intervention depending upon the needs of the students. As we all know, children develop at different rates, and the levels of intervention required for students to make sufficient progress varies. Due to this variability, many have asked, is RTI an effective approach, and how many tiers of intervention are needed to foster student success? Marston addressed this question in a 2005 article that synthesized results from three separate studies conducted by Vaughn, O'Connor, and Tilly. Each study implemented a three tier method for intervention among groups of students. In all three studies, tier one, two, and three students made greater gains than the control groups receiving no interventions. In addition, the researchers determined that students who continue to struggle with the curriculum given tier three intervention should be considered for special education services. Marston concluded that "the Vaughn, O'Connor, and Tilly studies provide positive outcomes showing the efficacy of the three-tier approach to response to intervention" (p. 544).

Johnson, Mellard, and Byrd (2005) outlined benefits and drawbacks of the RTI approach as presented by Scruggs (2003), Fletcher and Denton (2003), and Semrud-Clikeman (2003). Johnson et al. commented that the commonality between all of the above studies are that each retain several factors including evaluation of psychological processes, intra-individual discrepancies, and the rule out of external factors as possible contributors. Based on the synthesis of this research, Johnson et al. concluded that recommendations to improve the process of

identifying students with learning disabilities need to be developed. Furthermore, the authors mentioned the importance of "ensuring that students with LD are not only properly identified but also receive appropriate and effective interventions" (p. 571).

Unlike Gerber and Johnson et al., Kavale (2005) strongly opposed RTI. In his article, Kavale discussed the ramifications of eliminating the IQ-discrepancy model from the diagnosis of specific learning disability (SLD). Kavale reiterated the definition of SLD, and he suggested that if RTI replaces the IQ-discrepancy model it will be impossible to distinguish between different categories of special needs, such as reading disability versus the "slow learner" versus SLD. The author also talked about the ambiguity of RTI and the inevitable variability among teachers to determine success versus resistance to interventions. The author asserted that RTI emphasizes reading, and therefore this model largely ignores other areas of deficit such as math, writing, language, and comprehension. Kavale argued that the RTI method, if implemented, "will create diagnostic chaos," (p. 560) and proposed that RTI be used in conjunction with the IQ-discrepancy model to accurately identify students with special needs.

Like Kavale, Sternberg (2006) recommended a more practical approach in determining learning potential using the theory of successful intelligence. This theory suggests that effective teaching requires capitalizing on students' strengths and compensating for their weaknesses, discovering a balance in students' practical and creative abilities, capitalizing on students' life goals, and utilizing a variety of assessment measures. Sternberg presented results of several studies examining how students performed when this learning theory was implemented within their educational settings. In every study, students performed better than the control group when their teachers utilized this theory. These findings support his theory of mental self government, which was previously discussed.

Berkeley, Bender, Peaster, and Saunders (2009), although not opposed to RTI, pointed out some of the drawbacks of RTI. The authors suggested that RTI "excludes unexpected learning failure as part of identifying SLD because the presences of average or above average cognitive ability would not be documented" (p. 85). Furthermore, they indicated many of the interventions that are implemented as part of the RTI process do not have significant sources for empirical validation, and educators would not be able to distinguish SLD from various other disabilities.

Beyond pointing out the drawbacks of RTI, Berkeley, Bender, Peaster, and Saunders (2009) provided information on the progress each U.S. state has made in the use of RTI within their school systems. According to the authors, only 15 states have fully developed and are utilizing RTI models. "Twenty-two states are in the developing phase, 10 states are providing guidance to schools and districts, and 3 states are not currently in the process of developing a model or providing guidance or the information provided was unclear" (p. 87). According to the authors, the majority of states within this country acknowledge the benefit of RTI and much of the country will be using this model in the future.

Al Otaiba and Fuchs (2006) discovered that several specific characteristics, including low IQ, should not be ruled out when determining a student's ability to respond to interventions. Their study was conducted as part of an earlier investigation examining the effects of early literacy intervention on reading acquisition. In the earlier study, the examiners provided 312 students with reading intervention in kindergarten, first grade, both kindergarten and first grade, or typical classroom instruction for both years. The kindergarten and first grade interventions included *Ladders to Literacy* and *Peer-Assisted Learning Strategies.* Following the intervention,

the subjects were categorized as non-responsive, sometimes responsive, or always responsive based on their progress.

Of the 312 students that participated in the earlier study, 104 participated in the current study. The examiners administered a variety of standardized measures to the 104 participants to determine each student's ability level in an attempt to identify the characteristics of non-responders. The results indicated that non-responders typically performed poorly on measures of vocabulary, rapid naming, and verbal memory. The non-responders also exhibited behavioral difficulties. Consequently, Al Otaiba and Fuchs surmised that, "it is premature to conclude that verbal ability and IQ are not related to responsiveness" (p. 426).

Al Otaiba and Fuchs' article evokes many questions that challenge the efficacy of their study. Students participated in the interventions three days per week for 5-20 minutes per session. Children typically require a significant amount of time devoted to an area of learning in order to make sufficient progress; therefore, this limited exposure may have affected the subjects' ability to respond to the interventions. Furthermore, the participants' previous academic experience is unknown. Children who received an additional year of phonics training in preschool may have retained previously learned skills that assisted with their progress in the reading intervention.

In terms of the interventions, Al Otaiba and Fuchs utilized two different interventions that are supported by research; however, utilizing additional types of reading interventions would increase the reliability of their claim that non-responders present with particular characteristics. Non-responders for one type of intervention may be responders for another type of intervention. The type of instruction provided in the two interventions used may not have matched every child's learning style. As George (2005) stated, "every student is unique and, to a degree,

deserves and requires special attention and adaptation of the learning experience to fit those unique needs, interests, abilities, and attitudes" (p. 189). Moreover, the size of each intervention group was not indicated. Typically, instructors are able to provide students with more individual attention and monitoring in small groups as opposed to large groups.

The standardized measures that Al Otaiba and Fuchs utilized also raise questions. The examiners administered several subtests from various standardized testing measures rather than a complete battery, which provides bits and pieces of information regarding a child's abilities and fails to capture the big picture. For example, Mather and Woodcock (2001) stated in the examiner's manual for the Woodcock-Johnson III Test of Cognitive Abilities, "Although tests are the basic administration components of the battery, clusters of tests provide the primary basis for interpretation. Cluster interpretation minimizes the danger of generalizing from the score for a single, narrow ability to a broad, multifaceted ability" (p. 11). Administering an entire battery from one or more standardized measure could have altered the outcome, yielding entirely different characteristics of non-responders, or perhaps, no definitive characteristics of non-responders.

Contrary to Kavale (2005) and Al Otaiba and Fuchs (2006), O'Malley et al. conducted a study in 2002 based on their hypothesis that children with and without discrepancies, and children with average IQs do not have "underlying differences in cognitive ability, response to intervention, and related factors," (p. 31) and they successfully proved their hypothesis. Their study compared groups of children that met these criteria by examining developmental differences in eight precursor reading skills and reading related skills. The outcome of this study identified no differences in five out of the eight skill groups, which suggests that these children,

to a large degree, do not develop differently in this area, and therefore both groups should be capable of responding to reading interventions.

The following year (2003), Stage, Abbott, Jenkins, and Berninger conducted a similar study to predict how children would respond to reading interventions based on verbal IQ, attention ratings, verbal IQ-Word reading discrepancy, and reading related language abilities. Stage and colleagues collected data based on the subjects' performance on specific subtests from standardized tests, and created a growth curve to analyze the data. Their findings indicated that the presence of discrepancies did not predict response to early reading intervention. Stage and colleagues went on to state:

> Many students do not have a large enough discrepancy in first grade to qualify for special education services. They must fail for 1 to 3 years, until their discrepancy reaches an arbitrary size. Had they received services in their first grade general education classroom, when the probability of remediation was greatest, they might not have developed a discrepancy. Such practice, which is legally but not scientifically supported, poses a professional dilemma for educators. (p. 31)

The research conducted by O'Malley and colleagues, as well as the research conducted by Stage and colleagues, supports the need for further research in this area. Results from both studies make a case that intelligence does not necessarily predict ability to respond to interventions, and scientific proof to support the current methods of identification is virtually non-existent.

Fletcher et al. (2004) created a lengthy article addressing this issue, in which they argue for the incorporation of RTI into the current method of identification. They emphasized that response to intervention is a significant breakthrough, and if implemented correctly, would change the way both regular education and special education functions. The "adoption of RTI shifts the focus from a set of test scores that have limited utility for improving interventions to approaches that design and guide instruction to accelerate progress" (Fletcher et al., 2004, p.

313). Using RTI would provide all students that have difficulty learning the curriculum through typical means, with the opportunity to participate in specific interventions to address their needs.

Although the intention of the RTI method is to provide more students with the opportunity to be successful in school, disadvantages do exist. Semrud-Clikeman (2005) reviewed some of the potential problems that RTI could pose. Some of the drawbacks of this approach identified by Semrud-Clikeman include the following:

1. The majority of studies on RTI have focused only on children in kindergarten through second grade.

2. General education teachers are not trained in making decisions regarding the use of appropriate interventions.

3. Other areas of deficit, such as attention deficit, would be difficult to identify with this approach.

4. Definitive assessment tools have not been determined.

5. Specialists would not be able to identify the students' strengths and weaknesses across multiple domains.

The RTI process outlined by Barnett et al. (2004) and Compton et al. (2006) is practical and supported by empirical evidence; however, as Semrud-Clikeman (2005) pointed out, the RTI method is not flawless. Implementing this approach could potentially be costly and time consuming, and not every school system has the means to deal with such burdens (Gerber, 2005).

Hoover and Patton (2008) addressed the role of special educators in the RTI model. In the 1960's the primary role of special education teachers was educating students with disabilities in the confines of a substantially separate setting with little to no exposure of the general education curriculum. In the 1970's, their role moved from this separate and distinct area towards a more

inclusive setting, where students were educated within the general education setting, and pulled out for remedial services. In the 1980's students with special needs began spending even more time within the general education setting, with a concentration on remedial services as well as teacher support. This movement continued to grow within the 1990's, when inclusion classrooms became available for students with learning disabilities, where they received services throughout the day. "Currently, as we progress through the first decade of the 2000's, students with disabilities will increasingly receive their education within multitiered instruction using response to intervention as the primary basis for decision making" (p. 199).

Special educators, however, are not the only profession to play a major role in RTI. Mesmer and Mesmer (2008) discussed the impact of RTI on reading teachers. The authors stated, "RTI will alter the work of reading teacher because more than 80% of students identified for special education struggle with literacy, and the law names 'reading teachers' as qualified participants in the RTI process" (p. 280). In their article, Mesmer and Mesmer walked the readers through an actual case where RTI was used in the area of reading. During this process, the reading teacher played an integral role in identifying the student's areas of weakness, implementing interventions, and progress monitoring. The RTI process ultimately led to an increase in reading achievement for this particular student. Mesmer and Mesmer emphasized, "Despite the challenges with RTI, we have seen this approach increase the quantity and quality of instruction for struggling readers" (p. 289). As RTI is implemented in more schools, we will see a large number of reading teachers providing interventions for students who demonstrate deficits in this area.

Dexter, Hughes, and Farmer (2008) discussed the implementation and outcome of specific RTI programs by analyzing 11 previously conducted studies, based on their belief that

"it is necessary to bridge accountability goals with the assessment of students' instructional needs and to create programmatic structures that individualize instruction for learners who are at-risk for school failure" (p. 3). The inclusion criteria for the studies were publication in a peer-reviewed journal or edited textbook, use of at least two tiers of intervention, and quantifiable measurement of students' outcomes (academic, behavioral, systematic).

Of the 11 programs Dexter, Hughes, and Farmer identified, six were large district/state implemented and five were researcher-implemented. Each study was classified as either problem-solving, where the interventions were individually customized to meet student needs, or standard protocol, where pre-selected interventions were used. After analyzing the data, the authors concluded that, "each study examining the impact of an RTI program on academic achievement or performance included results that showed some level of improvement, with the authors attributing the changes to the RTI approach that was used" (p. 7).

Although the results of this analysis are promising for RTI, the outcome should be interpreted with caution. Dexter, Hughes, and Farmer pointed out that limitations include weak study designs, outcomes associated with the interventions are limited, and large scale RTI methods were not intended to carry out technically sound research. The authors indicated that more research should be conducted focusing on higher level reading skills as well as other academic areas.

Stecker, Fuchs, and Fuchs (2008) asserted that "Although sound instruction is paramount to successful implementation of RTI, assessment data should drive decision making. Therefore, we argue that progress monitoring comprises one of the most critical features of successful RTI implementation" (p. 10). The authors suggested that when providing intervention for struggling students, whoever is providing the intervention should continuously assess the students, a

minimum of once each month, to determine the student's progression through the curriculum and continue making sound decisions regarding intervention changes and long-term goals.

Bergstrom (2008) argued that successfully adopting and implementing RTI goes far beyond progress monitoring and scientifically based interventions; it requires a comprehensive school wide system reform. Bergstrom asserted that in order for RTI to be successful, one of the most important pieces in the reform is professional development geared towards educating teachers about the process. Without the appropriate professional development, the execution of such a program would be impossible.

Danielson, Doolittle, and Bradley (2007) agreed with Bergstrom in that professional development is crucial to make RTI successful, stating "To help improve student performance, the critical features of high-quality professional development should be in place, including professional development structures, such as teacher networks and study groups" (p. 633). The authors believed that without an extensive focus on professional development, it is impossible for RTI to be an effective process.

Like Bergstrom, Carney and Stiefel (2008) presented RTI as a school wide reform as well. They discussed the use of RTI as a problem-solving approach. To implement this approach, school systems must create an instructional support team (IST) to identify students who are at risk for failure and provide those students with appropriate interventions. Once the students have been identified, Carney and Stifel suggested that the team guide students through the RTI process.

Although this process has been implemented in some school systems throughout the United States, Carney and Stifel (2008) pointed out that "While RTI has come to schools through federal legislation, it has been left to educators and researchers to interpret and investigate the

best means of operationalizing this intent to ensure that students difficulties do not stem from instructional deficiencies" (p. 61). In other words, the vague wording of the changes in IDEA leave it open to interpretation. It is up to individual school systems to make decisions on how to structure RTI and ensure students who are at risk are provided with opportunities to become successful.

Restori, Gresham, and Cook (2008) indicated that despite the RTI option, the majority of states and local districts continue to use the IQ-achievement discrepancy method. The authors discussed the implications of the continued use of this method. Restori, Gresham, and Cook argued that this model has made it difficult to identify students and provide interventions early. Using IQ-achievement discrepancy, students are typically not identified until the third or fourth grade, causing them to be excluded from interventions during critical years of development. There is also little scientific evidence to support the IQ-achievement discrepancy approach. When comparing students with and without discrepancies, there is no meaningful difference between the two groups; whereas when comparing non-responders and responders to interventions (RTI), critical differences between the two groups exist. Furthermore, the authors contended that the IQ-achievement discrepancy model is used in an inconsistent manner, and many students do not receive interventions due to low IQ and commensurate achievement (slow learners).

Finally, Restori, Gresham, and Cook indicated that intelligence tests are used inappropriately and not in the way the creators of the tests intended, stating "Popular intelligence tests (CAS, WISC-IV, KABC-II, W-J III COG) actually discourage the use of their overall scores (i.e., Full Scale IQ) and strongly urge the user to use their tests for the purposes of identifying processing strengths and weaknesses" (p. 71). In other words, they created these tests

for the purposes of identifying students' learning strengths and weaknesses so that their strengths may be capitalized on and weaknesses compensated for. Despite these arguments, many districts throughout the country continue to use the IQ-achievement discrepancy method.

Many researchers are currently developing and conducting studies to support the efficacy of RTI and advocate for public schools to utilize this method (Barnett, Daly, Jones, & Lentz, 2004; Bergstrom, 2008; Fletcher, Francis, Morris, & Lyon, 2005; Compton, Fuchs, Fuchs, & Bryant, 2006). In order for the RTI method to be successfully implemented within schools, the educational system will rely on strong leaders and highly qualified teachers; however, the effort needs to begin at the state level, because at this point in time employing the RTI method is left to the discretion of each individual state.

Glover, DiPerna, and Vaughn (2007) suggested that "Although models for implementing interventions and monitoring student performance have been applied within educational settings, more research on core components of service delivery systems is needed" (p. 523). Furthermore, they asserted that "Research is also needed to evaluate the most effective selection, sequencing, intensity, and duration for individualized interventions" (p. 523). Additional studies need to be completed to support RTI and identify specific designs that will guide school systems in this process to encourage the use of RTI, if not mandate its use, across the United States.

Reading Disabilities

Although IQ scores are used to determine various types of learning disabilities, this study focuses solely on reading. Many studies have been conducted to support the use of IQ testing to identify reading disabilities; however, many studies have also been conducted that provide evidence suggesting IQ scores are not important in identifying reading disabilities.

Machek and Nelson (2007) conducted a study to examine the perceptions of practicing school psychologists on how reading disabilities (RD) should be operationalized, and what exclusionary criteria should be used when making determinations for special education eligibility within the public schools. Machek and Nelson pointed out that

> Defining reading disabilities (RD) is a historically imperfect practice that is tied to the larger history of federal definitions of learning disabilities (LD). There has long been a disconnect between the conceptual and operational definitions of LD, as well as a long history of controversy regarding both (p. 147)

As stated previously, because of this controversy, the Individuals with Disabilities Education Improvement Act (IDEIA, 2004) has been reauthorized stating that the IQ-achievement discrepancy model is no longer required. Unfortunately, this law leaves the definition of LD open to interpretation by individual school states.

In order to gain a more comprehensive understanding of how school psychologists view the definition of LD, more specifically RD, Machek and Nelson administered a 25 item survey to 1,480 randomly selected members of the National Association of School Psychologists (NASP) who currently work within the public school setting. The final response rate was 37%, which, according to Machek and Nelson, is consistent with other national surveys.

The results of the survey revealed that psychologists believe the most important criteria for the diagnosis of RD is treatment validity/response to intervention (32.8%), followed by cutoff on phonemic awareness measures (16.9%), cognitive processing (13.8%), and IQ-achievement discrepancy (10.6%). In addition, 84% of the respondents reported that exclusionary criteria should be incorporated into the definition of RD.

Machek and Nelson compared their results with a previous study completed by Speece and Shekitka in 2002. Speece and Shekitka's survey revealed that, at that time, the majority of psychologists suggested that cognitive processes was the most important criteria (77.6%),

followed by the IQ-achievement discrepancy model (61.9%), which according to Machek and Nelson, is most likely a result of the changes made to IDEIA (2004).

This study was limited in terms of low response rate and members surveyed. The researchers only received surveys back from 37% of the population, which restricts generalizability, and the respondents were all members of NASP. Machek and Nelson concluded that, "Since IDEIA 2004 leaves the door open for states…to develop their own RD classification standards, particular methods will need to continue to establish credible evidence of its benefits and be able to promote such benefits to all involved" (p. 154).

Further research needs to be conducted to explore various approaches to identifying reading disabilities. Furthermore, psychologists are only one member of the team that is responsible for diagnosing RD, and other members of the team should be considered when investigating criteria for RD.

Hettleman (2003) developed an article addressing this issue based on his belief that "early reading problems are too often blamed on weaknesses in intelligence or family background rather than specific deficits like phonological processing, which school systems can prevent and remediate. Low expectations of low-income children with low IQs are reinforced" (p. 10). This assertion is consistent with Edwards (2007) and Fletcher-Janzen and Daniel (2005) discussed in the previous sections. The author argued the importance of identifying children who are at risk for learning disabilities early so that educators can implement interventions that will increase their educational achievement. Hettleman identified the flaws in the IQ-achievement discrepancy model and discussed preliminary ideas for a pilot project that would enable all children to succeed. His pilot project emphasizes devoting attention to individual interventions in early childhood with gradual doses of increased intensity as needed.

Hettleman based his article on information that he has collected throughout his experience as an educational analyst and advocate in the City School System in Baltimore, Maryland. The information he presented is valuable, provides guidance for public education, and is supported by sound references and evidence from the relevant literature. Hettleman identified a population of students that often goes unnoticed, and he has attempted to provide those students with a strong voice in his article. He identified areas that should be strengthened in order to prevent students from falling through the cracks in the educational system.

Although Hettleman presented a compelling argument for a needed change, the information he obtained is pertinent to one region. Broadening his research area would benefit a larger population within public education. Furthermore, Hettleman's pilot project appears to lack the details necessary to implement the plan effectively within schools. He suggested that students who do not respond to treatment should receive additional small group instruction; however, he neglected to provide the reader with procedures on how to implement this instruction realistically. Most classroom teachers do not have the time and resources required to provide students with intensive small group instruction, and financial constraints inhibit many school systems from obtaining the support staff that is necessary for additional interventions. Hettleman suggested that the "reading specialist should provide all or most supplemental small group and tutoring instruction" (p. 24). His statement is presumptuous in that he assumes that few children will be identified as at risk, and/or every school that is taking on this project is small enough for the reading specialist to complete such a task.

Similar to Hettleman, Hitchcock, Prater, and Dowrick (2004) held concerns regarding a decrease in literacy skills among school-age children. They pointed out that, "According to the National Research Council (1998), over 40% of fourth and eighth-grade students were not able

to read well enough to perform assignments at grade level" (p. 89). Consequently, Hitchcock et al. conducted a study to examine the effects of tutoring and video self-modeling on reading fluency and comprehension skills. The authors utilized four first grade students with IQ scores ranging from 84 to 105 as their subjects. Two of the subjects had a diagnosis of specific learning disability, one subject was diagnosed with developmental delay, and one subject was in the process of being referred for a special education evaluation due to difficulty with academics. The examiners obtained baseline data prior to implementing the intervention, collected data two times per week during the intervention, and collected follow-up data 1 and 6 months post intervention. Results indicated that each student made gains in age equivalent scores in reading fluency and reading comprehension with the intervention. Furthermore, the students were able to maintain their skills following the intervention.

The data collected during Hitchcock et al.'s study provides evidence that reading comprehension interventions can be effective across students with varying abilities. The wide range in IQ levels and varying diagnoses demonstrates that this intervention can increase achievement with a diverse population. However, this study is not void of shortcomings. The limited subject base and absence of a control group makes the argument for this intervention less persuasive. Further research on this approach is necessary to provide additional support for its efficacy. In addition, the nature of the video self-modeling task is not only time consuming, but requires one to one teacher/student ratios, which is unrealistic in most school systems. Many school systems do not have the luxury of excess support staff to take on such challenges. The task also requires the use of digital video cameras, wireless microphones, and computer software, which could be unaffordable due to controlled budgets in many school systems.

Swanson and Howard (2005) also contributed to the literature on using IQ scores to identify learning disabilities. The authors argued, "Several studies have suggested that traditional intelligence tests provide a poor estimate of general ability" (p. 18). Based on this assertion, Swanson and Howard conducted a study to determine if using dynamic assessments is a valid measure to classify reading disabilities.

Swanson and Howard's subject base consisted of 70 students whose mean ages ranged from 10.37 to 12.79 that were classified as poor readers, reading disabled, both reading and math disabled, or skilled readers. Poor readers were defined as "children with verbal intelligence scores below a standard score of 96, but above 70" (p. 20). The authors selected subtests from various batteries to create their dynamic assessment measure, and the examiners administered the assessment to the participants providing probes or cues at strategic points throughout the testing. Once the examiners completed the testing, they compared the subjects' performance on the standardized measures to their performance on the dynamic assessment to determine whether the dynamic assessment provided enough information to detect a disability. The results indicated that using dynamic assessments yields valid measures in identifying children with disabilities. Swanson and Howard indicated that their results should be considered preliminary.

The findings in this study provide valuable information for practitioners and administrators in the field who are grappling with the decision to implement the newly developed method of identifying students with learning disabilities known as response to intervention. "The reauthorization of the Individuals with Disabilities Education Act [IDEA] now allows states the option of discontinuing the use of IQ-achievement discrepancy procedures as part of the LD identification process in favor of a response to intervention (RTI) approach" (Compton, Fuchs, Fuchs, & Bryant, 2006, p. 394). According to the U.S. Department of Education, IDEA serves

approximately 6.8 million students (U.S. Department of Education, 2006, ¶ 2); therefore, the implications of research such as Swanson and Howard's will affect a significant number of individuals within the near future.

Dynamic assessments would be an ideal tool to use with the response to intervention method because they allow examiners to, not only determine a child's ability level, but to also establish a child's learning potential by providing cues throughout the testing. Using these assessments, children would be able to demonstrate their ability to respond to interventions regardless of their IQ and achievement levels. In Swanson and Howard's study, 78% of the poor readers (children with IQs between 70 and 96) responded to the interventions provided and maintained the learned skills. Their findings support the idea that slow learners are able to respond to reading interventions regardless of their IQ scores.

Although Swanson and Howard's study provides evidence that children with low IQs are able to respond to appropriate interventions, an important factor in this study that may have influenced the outcome is their participants. Swanson and Howard's subjects were each receiving special education services at the time, which could have impacted their performance on the testing. In order to validate the use of dynamic assessments further, researchers should conduct studies using a more diverse subject base that consists of both students receiving services as well as students who are not receiving additional academic assistance.

Similarly to Swanson and Howard (2005), Joshi (2003) developed an article based on his theory that "the relationship between IQ and reading skills is not straightforward, and the association between the two has remained a controversial issue" (p. 248). In this article, Joshi introduced an alternative model for identifying poor readers, known as the componential model of reading. Unlike the IQ-achievement discrepancy model, the componential model provides

educators with recommendations regarding remediation for poor readers regardless of whether or not the students are identified as "reading disabled." The model emphasizes that reading is comprised of decoding, fluency, and comprehension, and if a student has a deficit in one of the three components, implementing specific remedial techniques will improve the deficit area. Therefore, Joshi believed that "children with poor reading comprehension skills can be trained to apply comprehension strategies when they read" (p. 260). The IQ-achievement discrepancy model identifies children as either "reading disabled" or "slow learners." The componential model identifies the area of reading a child is struggling with and provides a starting point for intervention.

Joshi's model for identifying students with disabilities is in its infancy. Although he has developed this model based on previous research that provides theoretical support, the model has not been scientifically tested. Researchers need to conduct follow-up studies to determine this model's efficacy and validity before the educational system utilizes this approach.

In addition, Joshi based the componential model on the belief that listening comprehension predicts reading abilities, which is supported by empirical evidence. However, this raises the question of who should administer the assessments, special needs teachers, speech-language pathologists, classroom teachers, or a collaboration of all educational professionals. Special needs teachers and classroom teachers typically assess reading abilities, and speech language pathologists typically assess listening comprehension. Furthermore, once the administrator is determined, the question of who will implement remediation for students that are identified as having deficits is raised. These questions need to be answered before educators are able to effectively utilize this model.

Contrary to Swanson and Howard (2005) and Joshi (2003), Swanson, Howard, and Saez (2006) claimed that "several recent studies have shown that differences between less skilled readers and skilled readers on measures of cognitive function are related to limitations in working memory" (p. 252). Because of these recent findings, Swanson, Howard, and Saez conducted a study to determine which components of working memory underlie reading comprehension and word recognition difficulties in less skilled readers.

During this study, 66 students were classified into one of four reading groups based on scores from standardized reading testing: children with RD, children with comprehension only deficits, skilled readers, and poor readers. The researchers then assessed each group in the following domains: real word and pseudo-word reading, word recognition, reading comprehension, fluid intelligence, processing speed, coding speed, phonological processing, verbal intelligence, short-term memory, working memory, and math achievement. The authors compared the results across the four groups and discovered variations in working memory among the groups. The skilled readers outperformed each group on measures related to working memory, children with comprehension only deficits performed better on working memory tasks than children with RD, and children with RD outperformed poor readers on working memory tasks. According to Swanson, Howard, and Saez, "these findings support the notion that different components of WM [working memory] underlie variations in reading performance" (p. 259).

Although the authors made some important discoveries as a result of their study, they indicated that more research needs to be conducted to examine the interaction between the executive and phonological systems during the act of reading. As mentioned previously, many scholars believe that IQ should not be the sole, or even the most important characteristic when identifying learning disabilities; however, these findings suggest that different aspects of IQ,

such as working memory, do play an important role in how a child is able to acquire reading skills.

Jenkins, Zumeta, Dupree, and Johnson conducted a study in 2005 on measuring gains in reading ability with passage comprehension. The authors discussed passage reading fluency (PRF) in relation to the RTI model. The participants in this study included 33 private school students between grades 2-10 with varying reading disabilities. The examiners assessed PRF for each participant at 3 different points in time. The results indicated that repeating passages did not inflate progress scores and showed significant variability on PRF measures among each testing point.

Although Jenkins et al. presented useful information regarding the assessment of reading fluency, weaknesses in their study are evident. Each participant in this study already had a diagnosis of reading disability. Therefore, the study does not provide readers with practical applications of identifying LD, which is the purpose of the RTI model. In addition, the authors did not discuss the types of intervention that each student received between assessment points or the variability in PRF scores among grade levels.

Their study leaves readers with several questions. Did each student receive the same type of intervention? If they did not, did the interventions they received impact their progress on PRF? Did the vast differences in grade level influence the subjects' performance? Is the amount of time between each assessment significant?

Cutting and Scarborough (2006) also investigated reading comprehension. Reading comprehension is a complex process, and many factors contribute to a child's ability to comprehend, including reading speed, word recognition, decoding, oral language, rapid serial naming and verbal memory. However, Cutting and Scarborough indicated that variability among

how educators measure comprehension may also contribute to perceptions of reading comprehension skills among students. Therefore, Cutting and Scarborough investigated the differences among several reading comprehension measures.

The researchers used 97 participants between the ages of 7.0 and 15.9 years with IQs of 80 or higher. The researchers administered three reading comprehension tests and a variety of measures of attention, verbal working memory, word oral language, recognition/decoding, reading speed, IQ, and serial naming. The researchers then compared the various measures and determined that word recognition/decoding and oral language skills made unique contributions to reading comprehension prediction, which is consistent with previous research. However, the contributions of word recognition/decoding varied considerably across comprehension measures, with nearly twice as much variance on one measure over another. "These findings suggest that different measures of reading comprehension may make differential demands on vocabulary knowledge and sentence-processing abilities" (p. 293).

Cutting and Scarborough's research findings have significant implications for the future of reading comprehension testing. Their demonstration of significant variability across measures suggests that educators should use a variety of measures when assessing reading comprehension. A single measure may provide educators with a skewed view of a student's ability to comprehend depending upon the demands of the assessment. This supports Swanson and Howard's (2005) assertion that using dynamic assessment measures yield a more accurate picture of a student's true abilities.

Svensson and Jacobson (2006) investigated another realm of reading. They conducted a longitudinal study to investigate the persistence of phonological skills from grades 3 through 12, and to examine the effectiveness of various definitions of dyslexia. The authors argued that there

are no specific criteria to define dyslexia, and "the long-standing apprehension concerning the definition of dyslexia is due to the discrepancy hypothesis, in other words, reading and spelling difficulties cannot be explained as a result of poor intelligence" (p. 6). Furthermore, Svensson and Jacobson pointed out that "the use of IQ might exclude some individuals that would otherwise have been regarded as dyslexic" (p. 6). The authors indicated that this has been an ongoing problem in education without resolution.

The subjects included 70 students who were divided into a poor reader group (40) and a carefully matched comparison group (30) that were followed over a 10 year period. The subjects were tested in the areas of word reading, non-word reading, phonological choice, figure relations, vocabulary, verbal ability and non-verbal ability in grades 2, 3, and 12. Once all of the data were obtained, the researchers compared each participant's performance over time. Based on the results, the researchers confirmed that phonological deficiencies that were evident in early school years did persist into adulthood and "a phonologically based definition of dyslexia exhibited more stability over time compared to definitions based on word-decoding and/or definitions based on IQ" (p. 13).

Johnston and Morrison (2007) conducted a study to determine if students with high IQs and students with low IQs differ in reading performance patterns. Participants included 109 poor readers ages 10 to 11 who were attending a specialized remedial reading unit outside of school. The students' reading ability was measured using the British Ability Scales Word Reading Test, and the students' IQ was tested using the short form of the Wechsler Intelligence Scale for Children Revised. The participants were matched with 105 same-age controls. Once matched, the students were split into two groups based on IQ scores 100 and below and 101 and above. All participants performed various reading tasks, including: regularity tasks to assess phonological

approach, a non-word reading task to measure phonological reading skills, and a phoneme deletion task to measure phonemic awareness skills. Results indicated that "High IQ poor readers showed deficits in both areas for reading age and even read non-words less well than the lower IQ poor readers" (p. 75) Furthermore, "They were also found to show impairments in reading regular words compared with both their reading-age controls and the low-IQ poor readers" (p. 75). This study provides evidence that students with lower IQs are not necessarily more impaired than students with higher IQ scores. In this study, the low IQ poor readers had less severe phonological deficits than high IQ poor readers, which provides support for their potential to improve their reading skills.

Although Johnston and Morrison provided strong support for students with low IQ scores, their study evokes an important question. Did the authors' use of the short form of the Wechsler Intelligence Scale for Children rather than the entire battery impact the validity of their IQ scores? As stated earlier, Mather and Woodcock (2001) indicated in the examiner's manual for the Woodcock-Johnson III Test of Cognitive Abilities, "Although tests are the basic administration components of the battery, clusters of tests provide the primary basis for interpretation. Cluster interpretation minimizes the danger of generalizing from the score for a single, narrow ability to a broad, multifaceted ability" (p. 11). Therefore, a full battery would provide more valid scores.

Ingesson (2005) studied the long-term effects of IQ in adolescents and adults with dyslexia by seeking an answer to the question "Is there a specific setback in cognitive ability, as measured with current IQ tests, in young people with developmental dyslexia as compared to normal subjects, as they grow up?" (p. 82). The author discussed the "Matthew Effect," coined by Stanovich (1986), which found that, "slow reading acquisition has cognitive, behavioural, and

motivational consequences that in turn slow the development of other cognitive skills and inhibit performance on many academic tests" (p. 82). Ingesson used the example of vocabulary. Poor readers are naturally unable to build their vocabulary as quickly and efficiently as their peers simply because they are not able to read as quickly and efficiently. This consequently may have a direct effect on their verbal IQ scores.

In this study, Ingesson used 65 participants were tested using the WISC-III at 12-years-old on average and retested after a mean interval of 6.5 years when subjects were between 14 and 25-years-old. Results showed a significant decrease between the initial and final assessments in verbal IQ. However, there was a significant increase in performance IQ among the two measures. In conclusion, Ingesson cautioned that educators should be careful when "making any decisions on the basis of a single intelligence test, and dyslexic children might be at risk to lag behind their peers in terms of VIQ, especially if they are not provided with suitable special education" (p. 81). Ingesson made a compelling argument for the caution of using a single measure in making educational decisions; however, the results of this study may have been impacted by many variables. It is possible that the time between the first and second IQ tests played a significant role in the outcome. In the interim, subjects may have had very different educational experiences. Some may have received intensive reading interventions, while other may not have, which could have impacted the follow-up IQ scores.

Palmer and Kevan (2007) also compared dyslexic readers to normal readers. In their study, Palmer and Kevan matched 44 subjects on nonverbal IQ, and administered a series of visual tasks and reading tasks to determine if visual sensitivity accounts for reading skills even when IQ is considered. Their findings indicated that nonverbal IQ was the best predictor of reading ability; however, they also found that "visual sensitivity explains an additional,

independent portion of the variance associated with reading ability" (p. 46). Unlike Ingesson, Palmer and Kevan did not take verbal IQ into consideration, which may have impacted the outcome of the study. Results may have been different if verbal IQ or full scale IQ were considered.

Conclusion

The studies that were reviewed are valuable; however, the results are contradictory. Some researchers have discovered that IQ is critical in identifying learning disabilities, while others argued to discontinue using IQ in identification methods. Some researchers have developed studies supporting the continued use of the IQ- achievement discrepancy model, while others argue for the use of the RTI model. These contradictions provide evidence that further research needs to be conducted in this area of education in order to propose sound solutions for this unresolved dilemma.

Although the findings from the studies presented in the literature review are contradictory, many of the studies do provide evidence that (a) children who have been labeled "slow learners" have been able to respond to interventions in the area of reading when the interventions were appropriate for their learning needs, (b) the IQ-discrepancy model is not the most effective method for identifying students with learning disabilities, and (c) RTI is a promising approach to identifying learning disabilities as it gives every child a chance to learn. Much of the literature presents empirical evidence with theoretical support; however, certain aspects of each study raise questions that challenge efficacy and reliability. Therefore, additional research needs to be conducted to build on the existing evidence to create a stronger argument for these children and their ability to learn with appropriate learning interventions. Without additional research, the argument remains weak and inconclusive. Conducting further research on the effects of response to intervention will allow the educational system to make sound resolutions on the best way to identify students with special needs and provide educational systems with the impetus to adopt this method.

The methodology of this study will be discussed in section 3, including the research design, specific research questions, setting/sample, a description of the intervention that will be provided to participants in the experimental group, the study process and data collection, instrumentation and materials, data collection and analysis including the hypothesis to be tested, the role of the researcher, and measures taken for ethical protection.

Section 3: Methodology

Introduction

This quasi-experimental quantitative study, using a pretest and posttest control group design, investigated the impact of a research based reading intervention on reading fluency, rate, and accuracy skills among fourth-grade students who were at risk for reading failure, as evidenced by the students' DIBELS Oral Reading Fluency scores. The design was chosen due to the nature of the research questions. According to Creswell (2003), "If the problem is identifying factors that influence an outcome, the utility of an intervention, or understanding the best predictors of outcomes, then a quantitative approach is best" (p. 21). Creswell also indicated that "In quasi-experiments, the investigator uses control and experimental groups" (p. 167). The pre and posttests allowed the researcher to determine each group's progress in reading fluency, rate, and accuracy to draw conclusions about the effect of the intervention.

The following elements of the study will be discussed in this section: the specific research question, setting and sample, a description of the intervention that was provided to participants in the experimental group, the study process and data collection, instrumentation and materials, reliability and validity, collection and analysis including the hypotheses to be tested, the role of the researcher, and measures taken for ethical protection.

Specific Research Questions

1. Is there a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group?
2. Is there a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group?

3. Is there a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group?

Hypotheses

*Ho*1: There is no significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*1: There is a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ho*2: There is no significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*2: There is a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ho*3: There is no significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*3: There is a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

The research question and hypotheses will be discussed in depth in section 3.

Setting and Sample

The setting of this study was an elementary school in a large urban district in the northeastern region of the United States. This study used a convenience sample to identify potential participants. A convenience sample, where participants "are chosen based on their convenience and availability" (Creswell, 2003, p. 156), was used because the researcher needed to identify participants based on current data, more specifically, the most current fourth-grade DIBELS Oral Reading Fluency scores. The sample consisted of fourth-grade students who have

placed in the at risk category based on their most recent DIBELS Oral Reading Fluency scores. Fourth-graders were chosen as the targeted population because research shows that many times, reading disabilities often do not present themselves and are not diagnosed until fourth-grade (Lipka, Lesaux, & Siegel, 2006). The NCES found that 40% of fourth graders did not meet the basic requirements set forth by the National Assessment of Educational Progress in Reading (NCES, 2006).

Twenty participants were chosen from the sample based on the following eligibility criteria: (a) students must have been in the fourth-grade, (b) students must have been in the at-risk category in reading based on DIBELS Oral Reading Fluency scores, (c) students must not have been identified as having a specific learning disability in the area of reading, and (d) students must not have been receiving supplemental reading instruction outside of the general education classroom. The decision to use 20 participants was derived through researcher consultation with the reading resource specialist (RRS) in the elementary school where the study took place. The RRS suggested using 20 participants based on the relatively small sample size according to current DIBELS Oral Reading Fluency data. The relatively small number of participants allowed the researcher to synthesize and analyze the data for each student in depth to draw sound conclusions about their response to the intervention provided.

Study Process and Data Collection

After the participants were identified, the researcher randomly dispersed the students into either the control group or the experimental group. Each participant was tested (pretest) using form A of the GORT-4 to assess reading fluency, rate, and accuracy. Following the pretest, the students in the control group receive reading instruction using the general education curriculum within the regular education classrooms, and the experimental group received supplementary

reading instruction, in addition to the general education reading instruction, using the Read Naturally reading intervention program outside of the regular education setting 4 days per week for 10 weeks. Following the 10-week period, students were retested using form B of the GORT-4 to assess reading fluency (posttest).

Threats to Validity

Although the study was conducted over 10 weeks period, absenteeism and ultimately limited exposure to the intervention may have caused the data to be inaccurate. Natural maturation during the 10-week period, previous exposure to reading interventions, and additional exposure to reading during the study could not be controlled and may have impacted the data. Additionally, the intervention that was used in the study has a subjective component, which may have caused the researcher to be inconsistent across subjects. Furthermore, inadequate statistical power due to the limited number of participants may have caused the conclusions drawn from the data to be inaccurate.

Intervention

Read Naturally (Software Edition)

Read Naturally is a research-based reading intervention program that focuses on developing reading fluency and promotes vocabulary development and reading comprehension. The software edition, which was used in this study, is a computer based program that enables student and teacher data to be stored securely on the computer. The program also "creates detailed reports that track student progress and indicate areas that need additional coaching" (Read Naturally, 2009).

Students begin this program by being placed at an appropriate reading level; from there, a reading goal is assigned to each student. After this goal has been established, the students follow

the progression of the program until their reading goal is met. The sequence of the program is as follows: Students select a story from their reading level; each reading level contains 24 story options. Then, students are presented with a series of key words that they click on to hear and find out the meaning of the word. Following the key words, the students write a prediction of what the story will be about based on the key words and the title. The students then move on to perform a cold read, which entails reading the story for 1 minute. When the minute is up, the timer sounds and the students click on the last word read. During the cold read, the students are instructed to click on any words that are difficult or that they are unable to read. After the cold read is completed, students read along with the story while the computer plays a recording of the story, then students practice reading the story without the recording until they are able to read it at the predetermined rate while the computer times each practice. Following the practice reads, students answer a series of comprehension questions and write about what they learned from the story. Finally, students read the story to the teacher to ensure that they are able to read the story at the goal rate. Once the students pass a story, they begin the process again with a new story. Validity and reliability will be discussed in further detail later in section 3.

Read Naturally was chosen as the reading intervention for this study because "the five critical components of reading instruction cited by the National Reading Panel, phonemic awareness, phonics, fluency, vocabulary and comprehension, may be found in varying degrees within the Read Naturally program" (Florida Center for Reading Research, 2009, p. 3). Although this program focuses primarily on fluency, all of the essential components of reading are woven in, creating a high level of success for students who struggle with reading.

Instrumentation and Materials

The instruments used in this study will be DIBELS and the GORT-4. The material used in the study will be the Read Naturally reading intervention program, which has been described in detail.

DIBELS

DIBELS is an assessment of early literacy development that focuses on phonological awareness, alphabetic principle, vocabulary, comprehension, and fluency with connected text that "allows educators to readily and reliably determine student progress" (DIBELS Data System, 2009, ¶ 2). However, for the purposes of this study, only the oral reading fluency portion will be used.

The oral reading fluency portion of DIBELS requires students to orally read a passage for one minute. The amount of words omitted and substituted, as well as hesitations of more than 3 seconds are marked as errors. If the student self-corrects within 3 seconds, the student receives credit. The oral reading fluency score is derived from the number of words correct per minute. DIBELS includes benchmark passages, which can be used as screening assessments throughout the school year as well as 20 alternate forms for progress monitoring (DIBELS Data System, 2009, ¶ 1). Validity and reliability will be discussed later in section 3.

Gray Oral Reading Test: Fourth Edition

The GORT-4 is a "norm-referenced, reliable, and valid test of oral reading rate, accuracy, fluency, and comprehension" (Wiederholt & Bryant, 2001, p. 3) for students between the ages of 6 years 0 months, through 18 years 11 months. One of the purposes of the GORT-4 is to "document students' progress in reading as a consequence of special intervention programs and

to serve as a measurement device in investigations where researchers are studying the reading abilities of school-aged students" (p. 4).

The test comes with two separate forms, form A and form B, each containing 14 individual reading passages. During the pretest, where form A was used, the student read a brief narrative and the examiner timed them and marks their miscues. The story was then removed from the student, and he or she answered a series of multiple choice comprehension questions that were read by the examiner. The examiner continued this process until the student received a fluency score of 2 or less and a comprehension score of 2 or less for a passage. Following the testing, the examiner scored the test and determined the standard scores for rate, accuracy, fluency, comprehension, and overall reading ability. For the purposes of this study, fluency, rate, and accuracy scores were used as pre and posttest measures. Comprehension scores were not used. Reliability and validity will be discussed in detail in the following section.

Reliability and Validity

Read Naturally

Read Naturally is based on "scientific research in fluency instruction" (Davidson, 2009, p. 6) that was developed in 1991 and has been evolving ever since. From 2002 through 2005, the company tested dozens of reading passages at each grade level by sending passages to schools throughout the country for the purposes of piloting with a diverse student population. Researchers compared words correct per minute (WCPM) on the passages with WCPM on benchmark passages and modified their reading passages based on the data. Results indicated that correlation coefficients ranged from .93 to .96, which is an indication of the program's validity. Additionally, in 2003, Read Naturally conducted studies on test re-test reliability for its fluency progress monitor within the same year, within consecutive years, and within every other

year and found that "across 121 coefficients, a point estimate is .907 with a 95% confidence interval from .899 to .915, which is well within acceptable psychometric standards" (Read Naturally Technical Information, 2002, p. 6).

DIBELS

According to the DIBELS technical report, "Even the least reliable DIBELS measure, Initial Sound Fluency (ISF), yields a reliability for .90 when administered 4 or 5 times for an approximate total of 5 minutes of assessment. Therefore, the reliability of 5 minutes of ISF would be .90." (DIBELS Data System, 2002, p. 9). A study conducted by Hintze, Ryan, and Stoner (2003) determined concurrent validity using a test that measures a similar construct known as the Comprehensive Test of Phonological Processes (CTOPP). The researchers indicated that "examination of the coefficients indicates that the DIBELS strongly correlates with most subtests and composite scores of the CTOPP" (p. 9).

Gray Oral Reading Test: Fourth Edition

The reliability of the GORT-4 was tested across three separate domains: content, time, and scorer. The testing manual provides statistics on all reliability coefficients which show that the test "evidences a high degree of reliability, which is consistently high across all three types of reliability, and rounds to or exceeds .90" (Wiederholt & Bryant, 2001, p. 69). According to the authors, "The magnitude of the reported coefficients strongly suggests that the test possesses little test error and that test users can have confidence in the results" (p. 70).

The GORT-4 testing manual provides information on content-description validity, criterion-prediction validity, and construct-identification validity. The rationale for the passages that are used in this test are based on current research, taking into account general interest, level of difficulty, and vocabulary. The authors also performed a series of tests to ensure that "the test

is nonbiased in regard to gender, race, and ethnicity" (p. 83). Furthermore, The GORT-4 correlates well with other tests that measure the same phenomenon. Based on research findings, the authors "have provided a considerable amount of evidence to show that the GORT-4 is a valid measure of reading performance and that it can be used with confidence" (p. 98).

Data Collection and Analysis

As discussed, data were collected prior to and immediately following the intervention. Once all data were collected, the researcher analyzed the data using an independent-measures *t* test. This method of analysis was chosen because the study compared a set of students taught by one method against a set of students taught by another method (Gravetter & Wallnau, 2005). The following hypotheses were tested:

*Ho*1: There is no significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*1: There is a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ho*2: There is no significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*2: There is a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ho*3: There is no significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*3: There is a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Participants' Rights

Role of Researcher

The researcher is employed as a speech-language pathologist by the school system where the study took place, and has been working in the elementary school where the study took place for the past 6 years. The researcher administered pre and posttests to all participants as well as implemented the reading intervention with participants in the experimental group.

Measures for Ethical Protection

The participants in this study and their legal guardians were provided with a detailed written description of the study including the purpose of the study, details regarding the reading intervention program, potential benefits, and the minimal potential risks. Legal guardians were provided with a letter of consent that they were required to sign if they chose to have their child participate in the study. The form contained contact information regarding where and when the researcher could be reached to field any questions the guardians or the participants may have had regarding the study. Legal guardians and their children's participation were voluntary, and they could have removed their children from the study at any time. Legal guardians and participants were also informed that personal information was not used in this study; each student's information was coded to maintain confidentiality, and students' names were changed as well. All student data were stored on the researcher's computer, which required a password to access. Copies of the assent and consent forms are included in the appendices.

Section 4: Results

Introduction

The purpose of this quasi-experimental quantitative study, using a pretest and posttest control group design, was to investigate the impact of a research based reading intervention on reading fluency skills among fourth grade students who are at risk for reading failure. An independent measures *t* test was used to measure the growth between pretest and posttest scores for the control group and the experimental group in order to answer the following three research questions: (a) Is there a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group? (b) Is there a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group? (c) Is there a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group?

The following elements of the study will be discussed in this section: overview of the study procedures, description of the sample, description of the participants, demographics of the experimental group and the control group, hypotheses, and statistical analysis for fluency, rate, and accuracy.

Overview of the Study

Twenty participants were chosen from a convenience sample based on the following eligibility criteria: (a) students must have been in the fourth-grade, (b) students must have been in the at-risk category in reading based on DIBELS Oral Reading Fluency scores, (c) students must not have been identified as having a specific learning disability in the area of reading, and

(d) students must not have received supplemental reading instruction outside of the general education classroom.

After the participants were chosen and parental consent forms were signed, the participants were randomly dispersed into either the experimental group or the control group. Each participant was tested (pretest) using form A of the GORT-4 to assess reading fluency. Following the pretest, the students in the control group received reading instruction using the general education curriculum in the regular education classrooms. The experimental group received supplementary reading instruction, in addition to the general education reading instruction, using the Read Naturally reading intervention program outside of the regular education setting 4 days per week for 10 weeks. Following the 10-week period, students were retested using form B of the GORT-4 to assess reading fluency (posttest). After all of the data were collected, The Statistical Package for the Social Sciences (SPSS) was used to conduct an independent measures *t* test to determine if there was a significant difference between the control group's growth according to pretest and posttest scores and the experimental group's growth according to pretest and posttest scores. When presenting the data, pseudonyms were used for all participants to maintain confidentiality.

Sample Description

The school where the research study took place had 61 fourth-grade students enrolled. Of the 61 fourth-grade students, 27 students were categorized as at risk for reading failure based on their DIBELS oral reading fluency scores. Three of the 27 students were excluded because they had learning disabilities in the area of reading and received special education services. This exclusion left 24 potential participants. Of the 24 potential participants, 3 were excluded due to lack of parental consent. This second exclusion left 21 potential participants. The first 20

participants who returned their parental consent forms were used in this study in order to have an even number of participants in each group.

Twelve of the participants were girls and 8 were boys. Of the 12 girl participants, 7 were Caucasian, 4 were African American, and 1 was Haitian. Of the 8 boy participants, 4 were Caucasian, 2 were African American, and 2 were Haitian. Nine students, 3 girls and 6 boys, were considered to be of low SES status based on free or reduced lunch. The participants ranged in age from 8.11 years-old to 10.1 years-old at the beginning of the study.

Control Group

Following a random selection process, the control group consisted of 5 girls and 5 boys. Three of the girls were Caucasian and 2 were African American. Two of the boys were Caucasian, 2 were African American, and 1 was Haitian. Two girls and 3 boys were considered to be of low SES based on free or reduced lunch.

Experimental Group

Following the random selection process, the experimental group consisted of 7 girls and 3 boys. Four of the girls were Caucasian, 2 were African American, and 1 was Haitian. Two of the boys were Caucasian and 1 was Haitian. One girl and 3 boys were considered to be of low SES based on free or reduced lunch.

Fluency Pretest and Posttest Results

Hypotheses

*Ho*1: There is no significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*1: There is a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Experimental Group

The experimental group sample size for the independent measures *t* test was $n = 10$. Each student's fluency pretest and posttest scores along with the difference between their fluency pretest and posttest scores are listed in table 1. Figure 1 shows the fluency pretest and posttest scores for each student in the experimental group, and Figure 3 shows the fluency growth based on pretest and posttest scores for each student in the experimental group. The results of the independent measures *t* test will be discussed in detail in this section.

Table 1

Experimental Group Pretest and Posttest Fluency Scores

Student	Pretest Score	Posttest Score	Difference
Brigit	6	11	+5
Terrance	8	12	+4
Thomas	11	12	+1
Kim	9	10	+1
Mary	5	9	+4
Justice	6	8	+2
Sarah	4	9	+5
Shanice	11	13	+2
Warren	10	12	+2
Aisha	6	10	+4

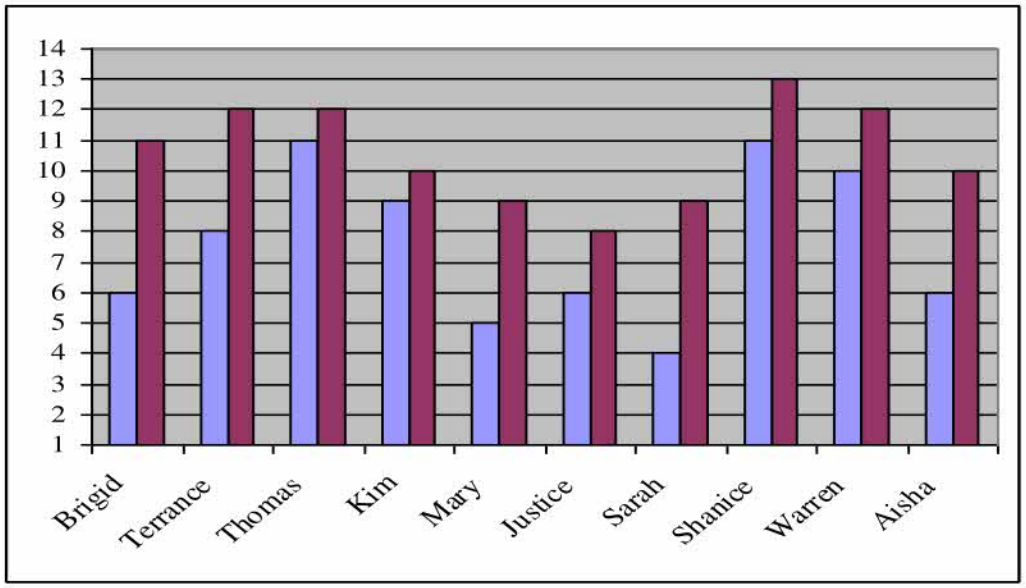


Figure 1: Experimental group pretest and posttest fluency scores

Control Group

The control group sample size for the independent measures *t* test was $n = 10$. Each student's fluency pretest and posttest scores along with the difference between their pretest and posttest scores are listed in table 2. Figure 2 shows the fluency pretest and posttest results for each student in the control group, and Figure 4 shows the fluency growth based on pretest and posttest scores for each student in the control group. The results of the independent measures *t* test will be discussed in detail in this section.

Table 2

Control Group Pretest and Posttest Fluency Scores

Student	Pretest Score	Posttest Score	Difference
April	4	6	+2
Isaiah	9	9	0
Michael	8	10	+2
Allison	5	9	+4
Jeffery	9	9	0
Malakai	12	13	+1
Nicholas	8	11	+3
Christina	10	11	+1
Serena	5	7	+2
Arissa	7	7	0

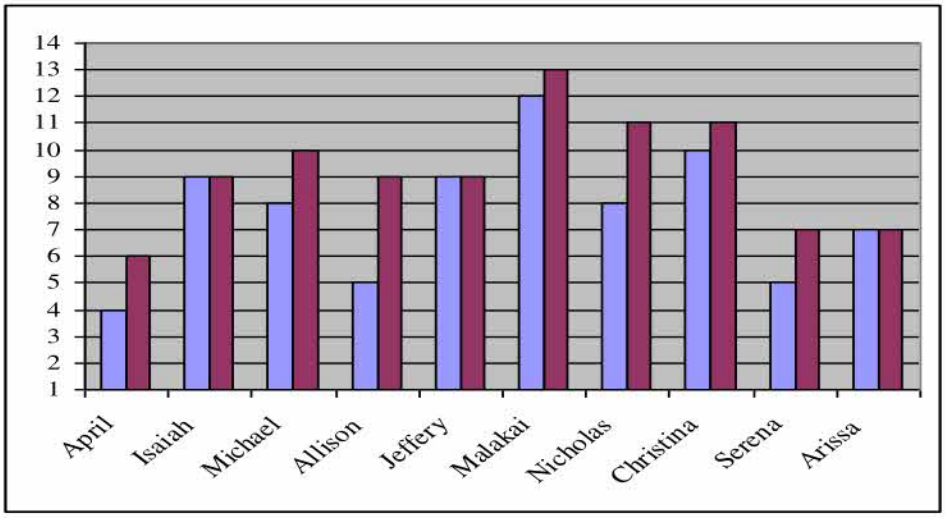


Figure 2: Control group pretest and posttest Fluency scores

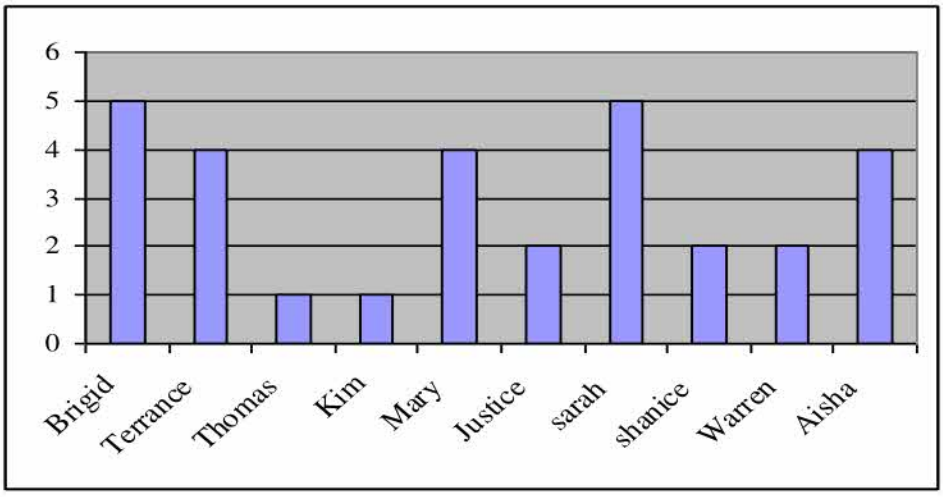


Figure 3: Experimental group fluency growth

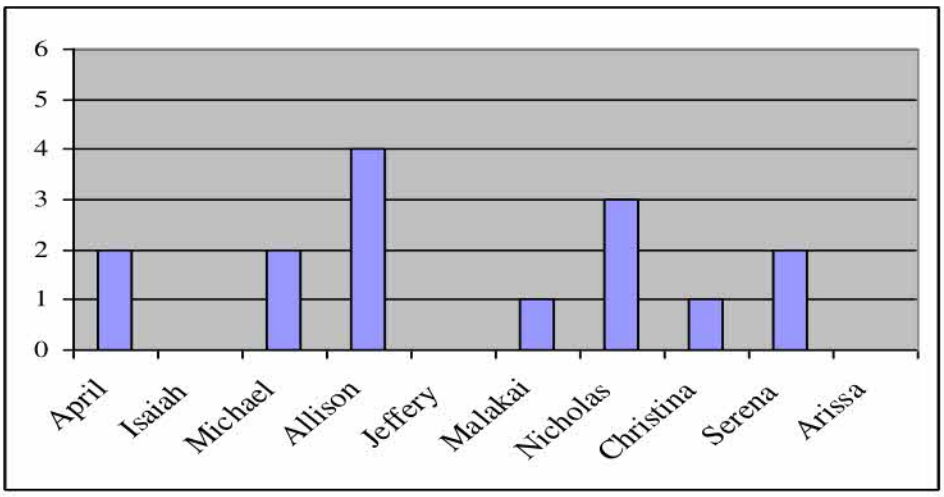


Figure 4: Control group fluency growth

Statistical Analysis of Fluency Data

Independent Measures *t* test

The independent measures *t* test was used to evaluate the mean difference between the experimental group's growth in the area of fluency based on pretest and posttest scores and the control group's growth in the area of fluency based on pretest and posttest scores. Because the independent measures *t* test only measures the treatment effect, the percentage of variance accounted for by the treatment was also calculated. According to Gravetter and Wallnau (2005), this measure examines whether the treatment causes the scores to vary. In addition, Hartley's F-

max test was also used to satisfy the homogeneity of variances assumption, which indicates that the two populations that are being compared have the same variances. Gravetter and Wallnau (2005) indicated that this is an important calculation because if the assumption is not met, the data cannot be interpreted meaningfully from an independent measures experiment.

Results

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.29 was obtained, which falls within the critical region. Because the *t* score value fell within the critical region, the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest scores of the experimental group is significantly higher than the growth based on pretest and posttest scores of the control group.

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 3.00 with a SD = 1.56 between pretest and posttest scores. The students in the control group had a M = 1.50 with a SD = 1.35 between pretest and posttest scores. This information is found on Table 3. This difference was significant, $t(18) = 2.29$, $p < .05$, $r^2 = .23$, which indicates that the 10-week reading intervention using the Read Naturally program had a medium effect on the students' reading fluency scores.

The F-max (1.55) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Table 3

Group Fluency Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Scores	Experimental	10	3.00	1.56	0.49
	Control	10	1.50	1.35	0.43

Table 4

Independent Measures t Test Fluency Data

				95% Confidence Level	
t	df	Mean Difference	Std. Error Difference	Lower	Upper
2.29	18	1.50	0.65	0.13	2.87

Rate Pretest and Posttest Results

Hypotheses

*Ho*2: There is no significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*2: There is a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Experimental Group

The experimental group sample size for the independent measures *t* test was $n = 10$. Each student's pretest and posttest scores for rate along with the difference between their pretest and posttest scores for rate are listed in Table 5. Figure 5 shows the rate pretest and posttest scores for each student in the experimental group, and Figure 7 shows the rate growth based on pretest

and posttest scores for each student in the experimental group. The results of the independent measures *t* test will be discussed in detail in this section.

Table 5

Experimental Group Pretest and Posttest Rate Scores

Student	Pretest Score	Posttest Score	Difference
Brigit	8	10	+2
Terrance	4	5	+1
Thomas	6	7	+1
Kim	9	10	+1
Mary	10	12	+2
Justice	7	7	0
Sarah	4	8	+4
Shanice	9	10	+1
Warren	7	9	+2
Aisha	7	10	+3

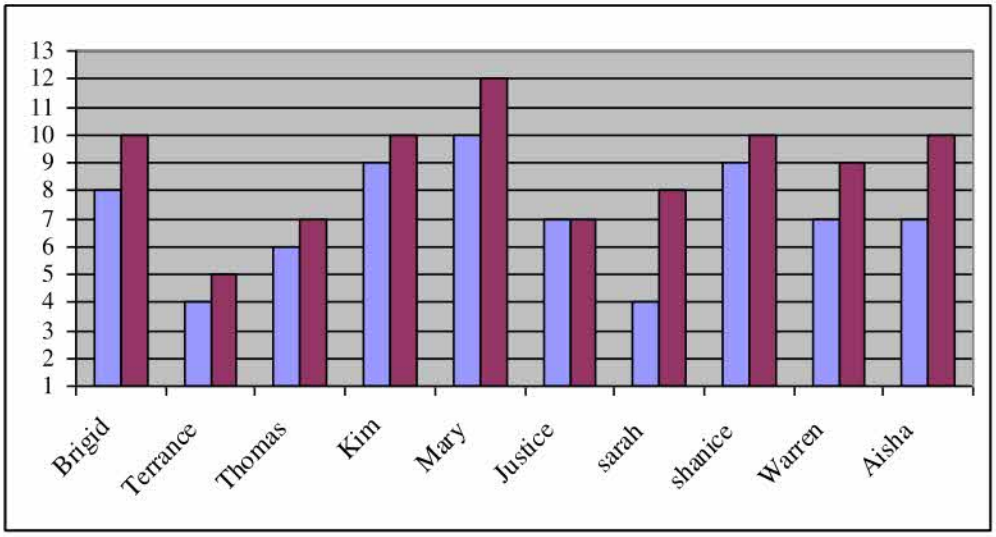


Figure 5: Experimental group pretest and posttest rate scores

Control Group

The control group sample size for the independent measures *t* test was $n = 10$. Each student's pretest and posttest scores for rate along with the difference between their pretest and posttest scores for rate are listed in Table 6. Figure 6 shows the rate pretest and posttest scores

for each student in the control group, and Figure 8 shows the rate growth based on pretest and posttest scores for each student in the control group. The results of the independent measures *t* test will be discussed in detail in this section.

Table 6

Control Group Pretest and Posttest Rate Scores

Student	Pretest Score	Posttest Score	Difference
April	5	6	+1
Isaiah	11	11	0
Michael	6	6	0
Allison	8	8	0
Jeffery	9	9	0
Malakai	10	10	0
Nicholas	7	8	+1
Christina	4	7	+3
Serena	7	7	0
Arissa	9	9	0

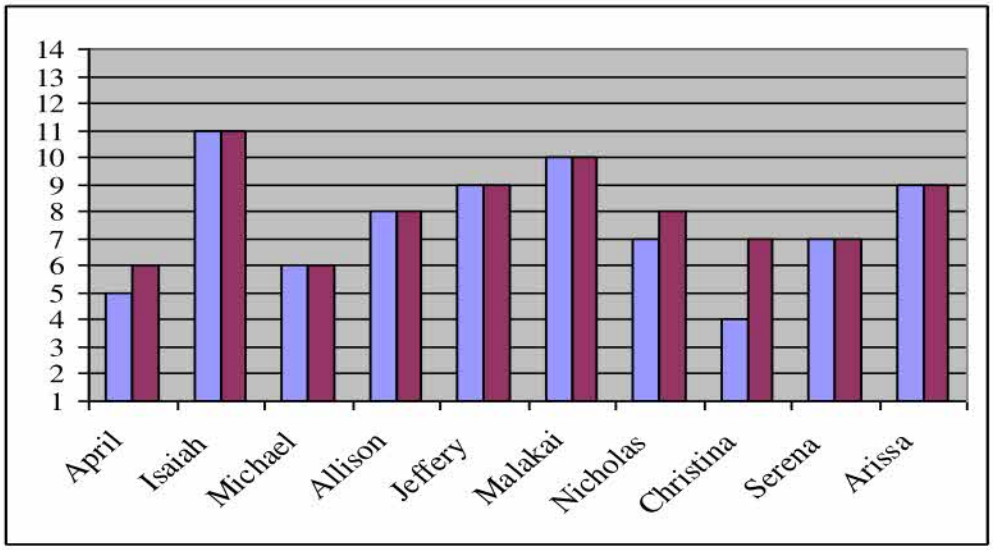


Figure 6: Control group pretest and posttest rate scores

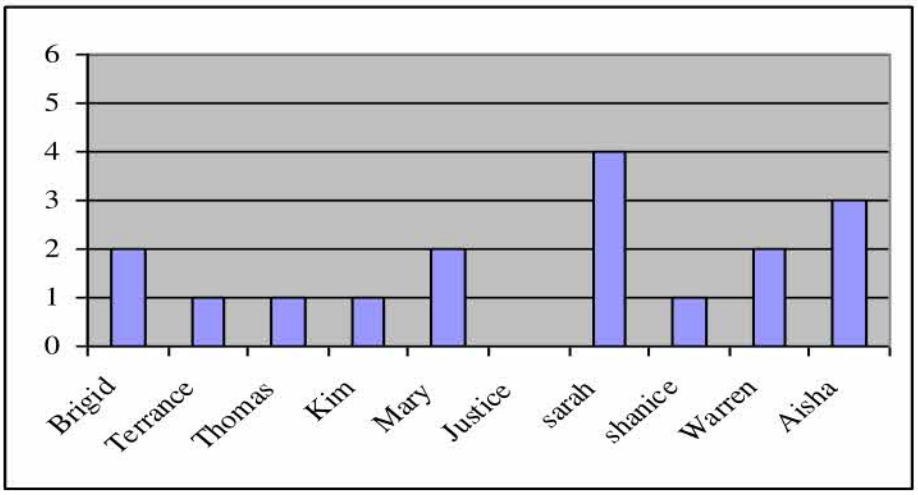


Figure 7: Experimental group rate growth

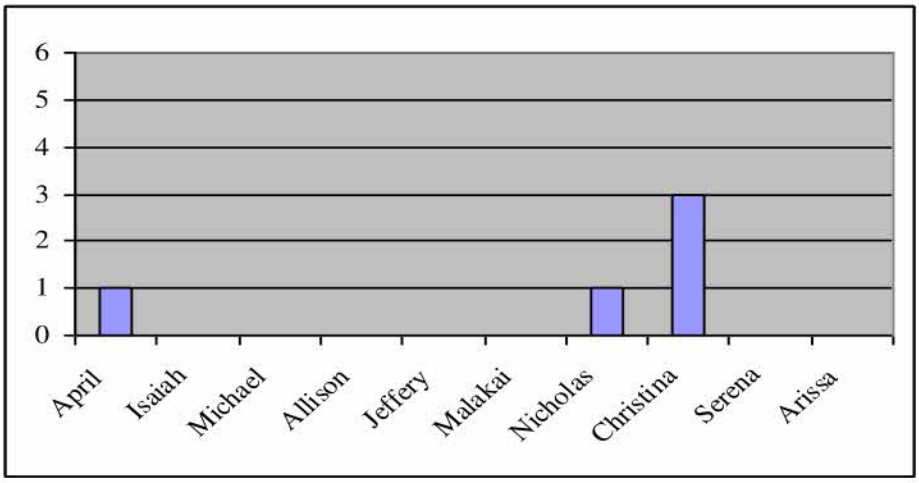


Figure 8: Control group rate growth

Statistical Analysis of Rate Data

Independent Measures t test

The independent measures *t* test was used to evaluate the mean difference between the experimental group's growth in the area of rate based on pretest and posttest scores and the control group's growth in the area of rate based on pretest and posttest scores. The percentage of variance accounted for by the treatment was also calculated, and Hartley's F-max test was used to satisfy the homogeneity of variances assumption.

Results

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.51 was obtained, which fell within the critical region. Because the *t* score value fell within the critical region, the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest rate scores of the experimental group is significantly higher than the growth based on pretest and posttest rate scores of the control group.

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 1.70 with a SD = 1.16 between pretest and posttest scores. The students in the control group had a M = 0.50 with a SD = 0.97 between pretest and posttest scores. This information is found on Table 7. This difference was significant, $t(18) = 2.51$, $p < .05$, $r^2 = .26$, which indicates that the 10-week reading intervention using the Read Naturally program had a large effect on students' reading rate scores.

The F-max (1.43) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Table 7

Group Rate Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Scores	Experimental	10	1.70	1.16	0.37
	Control	10	0.50	0.97	0.30

Table 8

Independent Measures t Test Rate Data

				95% Confidence Level	
t	df	Mean Difference	Std. Error Difference	Lower	Upper
2.51	18	1.20	0.48	0.20	2.21

Accuracy Pretest and Posttest Results

Hypotheses

*Ho*3: There is no significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

*Ha*3: There is a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Experimental Group

The experimental group sample size for the independent measures *t* test was $n = 10$. Each student's pretest and posttest scores for accuracy along with the difference between their pretest and posttest scores for accuracy are listed in Table 9. Figure 9 shows the accuracy pretest and posttest scores for each student in the experimental group, and Figure 11 shows the accuracy growth based on pretest and posttest scores for each student in the experimental group. The results of the independent measures *t* test will be discussed in detail in this section.

Table 9

Experimental Group Pretest and Posttest Accuracy Scores

Student	Pretest Score	Posttest Score	Difference
Brigit	9	12	+3
Terrance	9	10	+1
Thomas	7	12	+5
Kim	11	12	+1
Mary	12	14	+2
Justice	5	10	+5
Sarah	5	10	+5
Shanice	13	13	0
Warren	10	14	+4
Aisha	8	11	+3

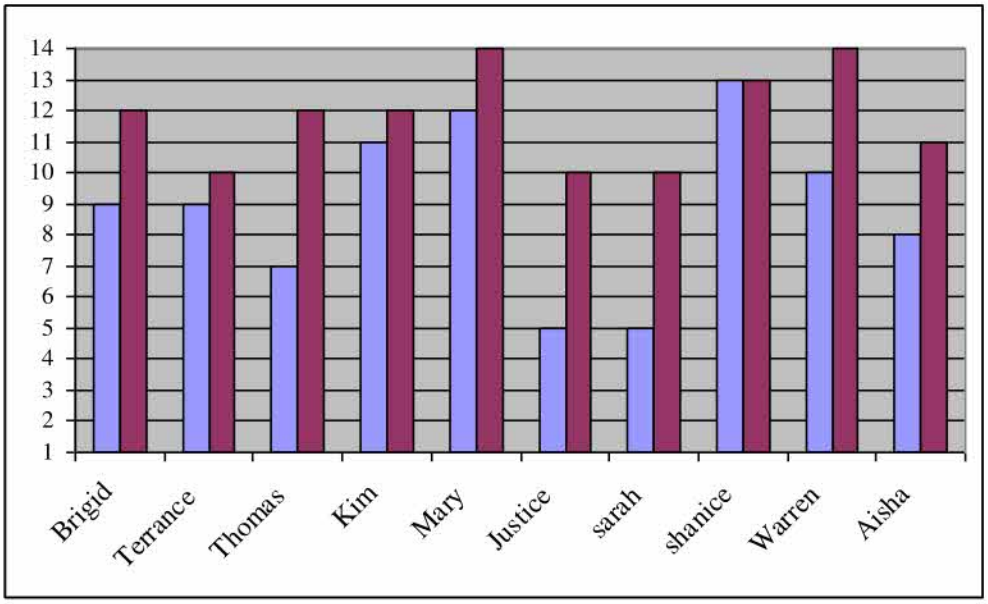


Figure 9: Experimental group pretest and posttest accuracy scores

Control Group

The control group sample size for the independent measures *t* test was $n = 10$. Each student's pretest and posttest scores for accuracy along with the difference between their pretest and posttest scores for accuracy are listed in Table 10. Figure 10 shows the accuracy pretest and posttest scores for each student in the control group, and Figure 12 shows the accuracy growth

based on pretest and posttest scores for each student in the control group. The results of the independent measures *t* test will be discussed in detail in this section.

Table 10

Control Group Pretest and Posttest Accuracy Scores

Student	Pretest Score	Posttest Score	Difference
April	6	8	+2
Isaiah	13	14	+1
Michael	8	9	+1
Allison	10	11	+1
Jeffery	11	11	0
Malakai	11	11	0
Nicholas	4	5	+1
Christina	9	10	+1
Serena	8	8	0
Arissa	9	12	+3

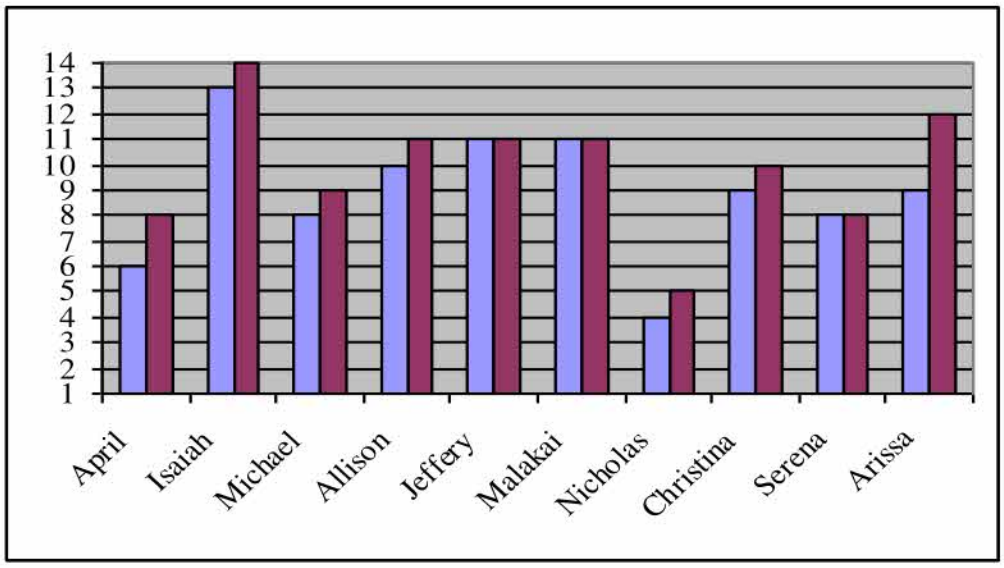


Figure 10: Control group pretest and posttest accuracy scores

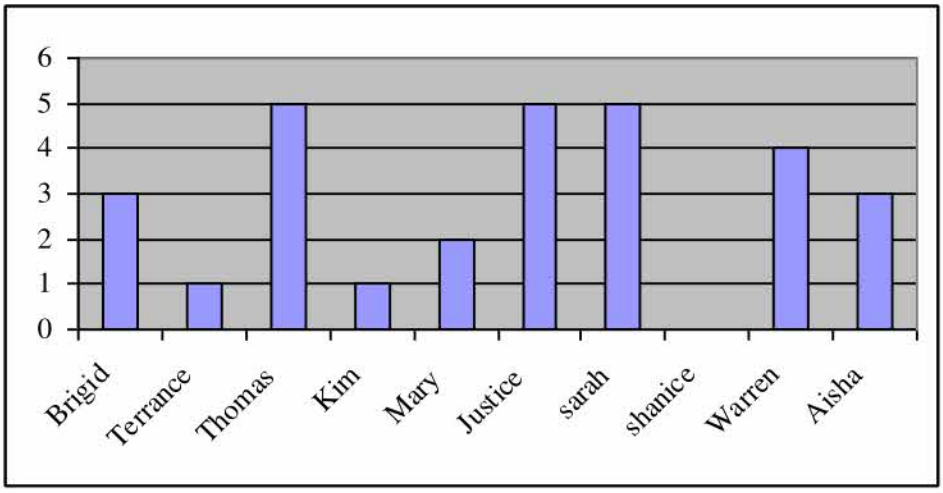


Figure 11: Experimental group accuracy growth

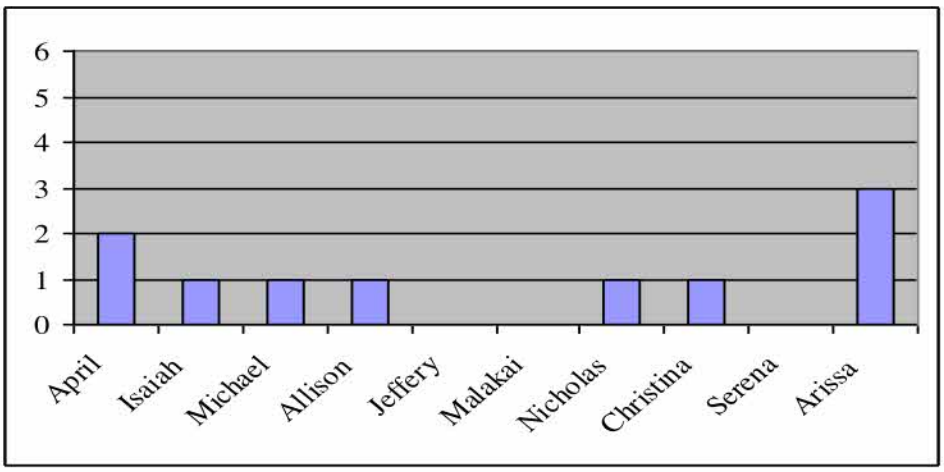


Figure 12: Control group accuracy growth

Statistical Analysis of Accuracy Data

Independent Measures *t* test

The independent measures *t* test was used to evaluate the mean difference between the experimental group's growth in the area of accuracy based on pretest and posttest scores and the control group's growth in the area of accuracy based on pretest and posttest scores. The percentage of variance accounted for by the treatment was also calculated, and Hartley's F-max test was used to satisfy the homogeneity of variances assumption.

Results

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.89 was obtained, which fell within the critical region. Because the *t* score value fell within the critical region the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest rate scores of the experimental group is significantly higher than the growth based on pretest and posttest rate scores of the control group.

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 2.90 with a SD = 1.85 between pretest and posttest scores. The students in the control group had a M = 1.00 with a SD = 0.94 between pretest and posttest scores. This information is found on Table 11. This difference was significant, $t(18) = 2.89$, $p < .05$, $r^2 = .32$, which indicates that the 10-week reading intervention using the Read Naturally program had a large effect on the students' reading rate scores.

The F-max (3.49) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Table 11

Group Accuracy Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Scores	Experimental	10	2.90	1.85	0.59
	Control	10	1.00	0.94	0.30

Table 12

Independent Measures t Test Accuracy Data

				95% Confidence Level	
t	df	Mean Difference	Std. Error Difference	Lower	Upper
2.89	18	1.90	0.66	0.52	3.28

Summary

According to the independent measures *t* tests that were conducted, the difference between the experimental group's growth and the control group's growth is significant in all three areas that were examined. Pretest and posttest results indicate that students in the experimental group significantly improved their fluency, rate, and accuracy following the 10-week reading intervention. Although the control group also improved in these areas, the improvement was not considered to be significant. The null hypotheses, which stated that there is no significant difference in the reading fluency, rate, and accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group, were rejected. Therefore, the answers to the research questions are yes: There is a significant difference in the reading fluency, rate, and accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group. These findings will be discussed in greater detail in section 5 along with conclusions that can be drawn from the study, limitations, recommendations, future research, and implications for social change.

Section 5: Summary, Conclusions, and Recommendations

Introduction

Using a quasi-experimental, pretest and posttest, control group, quantitative design, this research study investigated the impact of specialized instruction on reading fluency skills with students in Grade 4 who were identified as at risk for reading failure based on their DIBELS Oral Reading Fluency scores. According to the National Reading Panel, "Fluency is one of several critical factors necessary for reading comprehension, and despite its importance as a component of skilled reading, fluency is often neglected in the classroom" (National Reading Panel, 2009, ¶ 4). This study attempted to implement an RTI model for students who struggle with reading fluency to determine if this model could be successful in improving reading fluency skills. This research was conducted because many schools continue to use the IQ-achievement discrepancy model rather than adopting the RTI model. When schools strictly use the IQ-achievement discrepancy model, many students who struggle with reading fluency are not provided with the academic support they need to be successful, such as supplemental interventions, which causes these students to experience failure throughout their academic careers (Cooter & Cooter, 2004, p. 57).

This section will provide an overview of the study, summarize the research finding from section 4, and interpret the findings from section 4. Limitations, future research, and implications for social change will also be discussed.

Overview of the Study

The purpose of this quasi-experimental quantitative study, using a pretest and posttest control group design, was to investigate the impact of a research based reading intervention on reading fluency, rate, and accuracy skills among fourth-grade students who were at risk for reading failure. Twenty participants were chosen and randomly dispersed into either the quasi-experimental group or the control group. Each participant was tested (pretest) using form A of

the GORT-4 to assess reading fluency. Following the pretest, the students in the control group received reading instruction using the general education curriculum within the regular education classrooms, and the quasi-experimental group received supplementary reading instruction, in addition to the general education reading instruction, using the Read Naturally reading intervention program outside of the regular education setting 4 days per week for 10 weeks. Following the 10-week period, students were retested using form B of the GORT-4 to assess reading fluency (posttest). Once all of the data were collected, The Statistical Package for the Social Sciences (SPSS) was used to conduct an independent measures *t* test to determine if there was a significant difference between the control group's growth according to pre and posttest scores and the experimental group's growth according to pre and posttest scores.

Summary of Research Findings

This research was conducted to answer the three following questions: (a) Is there a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group? (b) Is there a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group? (c) Is there a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group?

The results of the statistical analysis using the independent measures *t* test revealed that there was a significant difference between the growth based on pretest and posttest scores for fluency, accuracy, and rate between the control group and the quasi-experimental group. According to the data, the research based intervention program Read Naturally enabled the students in the quasi-experimental group to improve reading fluency, rate, and accuracy skills through repeated practice and exposure.

The RTI approach used in this study followed the guidelines proposed by Barnett, Daly, Jones, and Lentz (2004), who suggested that educators identify children who are not learning the curriculum at the same rate as their peers and provide them with interventions based on their individual needs. The reading intervention received by the students in the quasi-experimental group was specifically geared towards targeting and improving deficit areas, including fluency, rate, and accuracy. Each component of reading that was examined in this study will be discussed.

Fluency

Following the 10-week reading intervention, the quasi-experimental group had a mean growth of 3.00 points between their pretest and posttest scores, and the control group had a mean growth of 1.50 points between their pretest and posttest scores. This difference was considered to be significant with a medium effect. Gravetter and Wallnau (2005) indicated that this measure of effect examines whether the treatment causes the scores to vary. According to the results, the 10-week reading intervention program had a medium effect on the students' reading fluency scores.

Considering the individual participants within each group, all of the students in the quasi-experimental group improved by at least 1 point, while three of the students in the control group did not make any improvement at all. Five of the students from the quasi-experimental group were able to make a gain of more than 2 points, while only two students from the control group were able to make a gain of more than 2 points.

The vast difference in growth between the groups suggests that the reading intervention played an integral role in the students' improvement. Mesmer and Mesmer (2008) conducted a similar study which also resulted in an increase in reading achievement. They emphasized "Despite the challenges with RTI, we have seen this approach increase the quantity and quality

of instruction for struggling readers" (p. 289). With targeted quality instruction, students were able to improve their reading skills in this study as well as in Mesmer and Mesmer's study.

These findings are also consistent with McComas et al.'s (2009) study in which second and third-grade students were able to significantly improve reading fluency skills following reading intervention which included incentives, modeling, repeated reading, and error correction. Following this intervention, which targeted reading fluency, students were able to significantly increase their scores. Huang, Nelson, and Nelson (2008) also reported an increase in reading fluency skills for two second-grade students who participated in a 10-week repeated reading intervention. They indicated that following the targeted reading intervention, both students made a significant increase in terms of words correct per minute, with a large effect according to their statistical analysis. These results are consistent with the results from this study.

Rate

In addition to reading fluency, reading rate was also examined in this study because it is so closely related to reading fluency. According to Mesmer (2010) the construct of fluency includes efficient reading rate. Furthermore, part of the overall goal of the Read Naturally intervention program is to improve reading rate. Following the 10-week reading intervention program, the quasi-experimental group had a mean growth of 1.70 points between their pretest and posttest scores, and the control group had a mean growth of 0.50 points between their pretest and posttest scores. This difference was considered to be significant with a large effect. Therefore, the 10-week reading intervention had a large impact on the students' reading rate scores.

Examining the performance of the individual participants in each group, nine of the students from the quasi-experimental group improved by at least 1 point, where only three of the

students from the control group improved by at least 1 point. Five of the students from the quasi-experimental group were able to make a gain of 2 or more points, where only one of the students from the control group was able to make a gain of 2 or more points.

Again, the sizable difference in growth between the groups suggests that the reading intervention played a vital role in the students' improvement, but this time with a large effect according to the data analysis. These findings are consistent with Dexter, Hughes, and Farmer's (2008) examination of various studies that implemented RTI. They stated that "each study examining the impact of an RTI program on academic achievement or performance included results that showed some level of improvement" (p. 7). The results of this study are also consistent with Musti-Rao, Hawkins, and Barkley's (2009) findings. In their study, 12 fourth-grade students significantly improved their oral reading rate following a repeated oral reading intervention program. The researchers compared two methods of reading practice, repeated silent reading versus repeated oral reading, and found that participants made significantly greater gains when participating in repeated oral readings. Read Naturally uses repeated oral reading as one of its key components, and students from this study experienced similar growth using this strategy.

Accuracy

In addition to reading fluency and reading rate, reading accuracy was also examined in this study as it, too, is closely related to reading fluency. Mesmer (2010) indicated that the construct of reading fluency includes accurate word reading in addition to efficient reading rate. Also, part of the overall goal of the Read Naturally intervention program is to improve reading accuracy. Following the 10-week reading intervention program, the quasi-experimental group had a mean growth of 2.90 points between their pretest and posttest scores, and the control group had a mean growth of 1.00 point between their pretest and posttest scores. This difference was

considered to be significant with a large effect, which again implies that the 10-week reading intervention had a large impact on the students' reading accuracy scores.

Examining the performance of the individual participants in each group, nine of the students from the quasi-experimental group improved by at least 1 point, where only seven of the students from the control group improved by at least 1 point. Seven of the students from the quasi-experimental group were able to make a gain of 2 or more points, where only two of the students from the control group were able to make a gain of 2 or more points. Again, the considerable difference in growth between the groups suggests that the reading intervention played a fundamental role in the students' improvement with a large effect according to the data analysis. The results of this study are consistent with findings from Vadasy, Sanders, and Tudor's (2007) study in which they investigated the impact of reading intervention on a group of 42 students in second and third-grade. Students in the experimental group received a total of 21 hours of targeted individualized reading instruction. Following the intervention, these students were able to make significant gains in their reading accuracy when compared with the control group.

These findings are also consistent with Knezek and Christensen's (2007) study, which examined the use of technology-based reading interventions across 25 public school districts. In their study, reading interventions were implemented with first and second-grade students in 18 school districts, and 7 school districts with matching demographics served as a control group. The authors indicated the major findings suggested the use of technology-based interventions fostered reading accuracy among first and second-grade students.

Limitations

The major goal of this study was to examine the impact of RTI on student achievement by determining if participating in reading intervention would improve students' reading fluency scores. Although the data indicate that the reading intervention did in fact improve students' reading fluency skills, this study had limitations that make it difficult to generalize the results to the general population. One important limitation was the relatively small sample size used in this study. The goal of any research study is to generalize the results from a selected sample back to the entire population (Gravetter & Wallnau, 2005). However, because this study used a sample of only 20 students, additional research corroborating the results would be beneficial in generalizing the results to the population.

Another limitation of this study was that it focused solely on reading fluency and its related components. In order to truly determine if RTI is effective, other areas of learning should be investigated. Moreover, this study was limited to fourth grade students. If this study had used participants from other grade levels the results may have been different. To validate the effects of RTI, interventions should be implemented and studied with students from various grade levels. Also, this study only used one reading intervention. The results may have been different if a variety of reading interventions were used. Consequently, it is necessary to investigate other reading programs to determine if RTI is effective using various interventions.

In addition to the limitations posed by the design of the study, the researcher also could not account for the behavior of the participants in this study. The participants had varying levels of motivation, and therefore they each demonstrated varying amounts of effort and dedication to improving their reading fluency skills. Some students were highly motivated and were extremely focused when participating in the intervention while other students were disinterested and simply

went through the motions. Also, it is unknown how much time students in the study practiced reading outside of the allotted time, which may have impacted some of the student's posttest scores.

Future Research

Due to the limitations in this study, further research needs to be conducted to investigate the effects of RTI. As stated earlier, research should be conducted using various interventions in all areas of the curriculum. This study focused solely on reading fluency, rate, and accuracy; however, it would be beneficial to examine phonics, reading comprehension, mathematics, and other academic skills to determine if RTI is effective in increasing student achievement in all learning areas.

In addition, future research should focus on examining RTI in all grade levels. This study used students in Grade 4. Investigating the outcome of implementing RTI across all grade levels would provide valuable information. Also, future research should focus on putting RTI into practice with larger samples of students. The small sample size in this study makes it difficult to generalize the results to the entire population. If larger samples are used, the results could be more reliable and researchers would be able to generalize the results. It may also be beneficial to replicate this study. Conducting this study using a new group of participants to determine if the results are consistent would provide support for the accuracy of the findings.

Recommendations for Action

Implementing these suggestions in future research could allow the educational system to make sound resolutions on the best way to increase student achievement and identify students with special needs. It will also provide educational systems with the impetus to adopt this method. Administrators in special education departments who determine qualifications for

specialized instruction should pay careful attention to research in this area to aid in making decisions regarding the use of RTI. Bergstrom (2008) argued that successfully adopting and implementing RTI goes far beyond progress monitoring and scientifically based interventions; it requires a comprehensive school wide system reform. According to Berkeley, Bender, Peaster, and Saunders (2009), only 15 states have fully developed and are utilizing RTI models. Therefore, the USDOE should look at the results of this study as well as future research to aid in determining the most effective method of identifying learning disabilities, and ultimately implementing new special education standards for all school systems.

The data from this study support RTI as an effective way to improve student achievement, and the results will be shared with the researcher's school system to provide department heads and decision makers with information on how to implement RTI successfully within each school. Educators who have students who are at risk for reading failure should also pay attention to the results of this study as well as future research to assist in determining methods of instruction and development of lesson plans and curriculum. It is recommended that the results from this study be shared with teachers and other faculty members through professional development to provide educators with a concrete example of how valuable this approach can be with struggling students.

In order for this method to be implemented throughout the system, educators must develop a deep understanding of and belief in this process. Furthermore, providing teachers and other faculty with the appropriate training will allow them to build confidence in their ability to use RTI with their students. Bergstrom (2008) asserted that in order for RTI to be successful, one of the most important pieces in the reform is professional development geared towards educating

teachers about the process. Without the appropriate professional development, the execution of such a program would be impossible.

Implications for Social Change

As stated in section one, the USDOE has created new guidelines for identifying learning disabilities that allow school systems to adopt the RTI method in lieu of or to use in conjunction with the IQ-achievement discrepancy model. Although RTI has been implemented in some school systems throughout the United States, Carney and Stifel (2008) pointed out that "While RTI has come to schools through federal legislation, it has been left to educators and researchers to interpret and investigate the best means of operationalizing this intent to ensure that students difficulties do not stem from instructional deficiencies" (p. 61). In other words, the vague wording of the changes in IDEA leaves it open to interpretation. It is up to individual school systems to make decisions on how to structure RTI and ensure students who are at risk are provided with opportunities to become successful. This study, along with future research, will allow the Department of Education, school administrators, and other educational leaders to evaluate the RTI process and establish a protocol for its use in schools across the country to improve student achievement, ultimately effecting social change.

Without relevant research on RTI, many school systems would continue to use the IQ-achievement discrepancy model and perpetuate the ongoing problem of students who do not qualify for special education services falling through the cracks. These students may not be given the chance to increase their achievement skills, even though they struggle with the curriculum simply because they do not qualify to receive specialized instruction (Ukrainetz, 2006). Implementing RTI in schools across the country will allow all students the opportunity to participate in interventions geared towards their specific learning needs. With this targeted

instruction, students may be able to improve very precise areas of deficit to advance their overall educational performance, which in turn will afford them greater opportunities in the future.

Just as everything in our society changes and evolves, the field of education must continuously undergo transformations in order to provide students with the knowledge needed to become productive members of our ever-changing world. The only way to achieve this evolutionary goal is to continue conducting research on the cutting edge issues in education. The identification methods debate has been ongoing over the past two decades, and continues to be unresolved. This research, along with future research, could help educational leaders to promote social change by coming to a consensus on how to best solve this longstanding problem so that we may move forward and continue to make advancements in the field of education, ultimately advancing our society.

REFERENCES

AIMSweb Progress Monitoring and RTI Systems. Retrieved from http://www.aimsweb.com/measures-2/maze-cbm/

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities, 39*, 414-431.

Barnett, D. W., Daly, E. J., Jones, K. M., & Lentz, F. E. (2004). Response to intervention: Empirically based special service decisions from single-case designs of increasing and decreasing intensity. *The Journal of Special Education*, *38*, 66-79.

Bergstrom, M. K. (2008). Professional development in response to intervention: Implementation of a model in a rural region. *Rural Special Education Quarterly, 27*(4), 27-36.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities, 42*(1), 85-95.

Brown-Chidsey, R., Johnson, P., & Fernstrom, R. (2005). Comparison of grade-level controlled and literature based maze CBM reading passages. *School Psychology Review, 34*(3), 387-394.

Carney, K. J., & Stiefel, G. S. (2008). Long-term results of problem solving approaches to response to intervention: Discussions and implications. *Learning Disabilities: A Contemporary Journal, 6*(2), 61-75.

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities, 36*, 151-164.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98,* 394-409.

Cooter, K. S., & Cooter, R. B. (2004). One size does not fit all: Slow learners in the reading classroom. *The Reading Teacher, 57*(7), 680-684.

Creswell, J. (2003). Research design: Qualitative, quantitative, and mixed methods approaches. Thousand Oaks, CA: Sage.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language, proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*(3), 277-299.

D'Angiulli, A., & Siegel, L. S. (2003). Cognitive functioning as measured by the WISC-R: Do children with learning disabilities have distinctive patterns of performance? *Journal of Learning Disabilities, 36*, 48-58.

Danielson, L., Doolittle, J., Bradley, R. (2007). Professional development, capacity building, and research needs: Critical issues for response to intervention implementation. *School Psychology Review, 36*(4), 632-637.

Dethorne, L. S., & Watkins, R. V. (2006). Language abilities and nonverbal IQ in children with language impairment: Inconsistency across measures. Clinical *Linguistics and Phonetics, 20*(9), 641-658.

Dexter, D. D., Hughes, C. A., & Farmer, T. W. (2008). Responsiveness to intervention: A review of field studies and implications for rural special education. *Rural Special Education Quarterly, 27*(4), 3-9.

DIBELS Data System. *Assessment Committee Analysis of Reading Assessment Measures*. (2009). Retrieved from http://dibels.uoregon.edu/techreports/dibels_5th_ed.pdf

DIBELS Data System Description and Sample Reports. Retrieved from https://dibels.uoregon.edu/samples/index.php

Edwards, O. W. (2006). An analysis of the differential impact of IQ: Considering consequential validity and disproportionality. *Research in the Schools, 14*(1) 29-39.

Edwards, O. W. (2007). Special education disproportionality and the influence of intelligence test selection. *Journal of Intellectual and Developmental Disability, 31*(4), 246-248.

Epelbaum, D. (2007). Multiple intelligence assessments give insight into reading comprehension difficulties and potential: A case study. *The International Journal of Learning, 14*(5), 243-251.

Eris, B. (2008). Intelligence: An analysis of the American experience from the perspective of critical theory. *Educational Sciences: Theory and Practice, 8*(1), 79-87.

Fer, S. (2005). Validity and reliability of the thinking styles inventory. *Educational Sciences:Theory and Practice, 5,* 55-68.

Fletcher, J. M., Coulter, W. A., Reschly, D. J., & Vaughn, S., (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia*, *54,* 304-329.

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34,* 506-522.

Fletcher-Janzen, E., & Ortiz, S.O. (2006). Cultural competence in the use of IQ testing with culturally and linguistically diverse children. *Gifted Education International, 21*, 137-150.

Flynn, J. R., & Weiss, L. G. (2007). American IQ gains from 1932 to 2002: The Wisc subtests and educational progress. *International Journal of Testing, 7*(2), 209-224.

Francis, D. J., Fletcher, J. M., Stuebing K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities*, *38*, 98-108.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disabilities Quarterly, 25*(1), 33-45.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness to intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18*, 157-171.

Gardner, H. & Moran, S. (2006). The science of multiple intelligences theory: A response to Lynn Waterhouse. *Educational Psychologist, 41*(4), 227-232.

Gerber, M. M. (2005). Teachers are still the test: Limitations of response to instruction strategies for identifying children with learning disabilities. *Journal of Learning Disabilities, 38*, 516-524.

Glover, T. A, DiPerna, J. C., & Vaughn, S. (2007). Introduction to the special series on service delivery systems for response to intervention: Considerations for research and practice. *School Psychology Review, 36*(4), 523-525.

Glutting, J. J, Watkins, M. W., Konold, T. R., & McDermott, P. A. (2006). Distinctions without a difference: The utility of observed versus latent factors from the WISC-IV in estimating reading and math achievement on the WIAT-II. *The Journal of Special Education, 40(*2), 103-114.

Gravetter, F. J., & Wallnau, L. B. (2005). Essentials of statistics for the behavioral sciences. 5th Edition. New York, NY: Thomson Publishing

Greenspan, S. (2006). Functional concepts in mental retardation: finding the natural essence of an artificial category. *Exceptionality, 14*(4), 205-224.

Hall, I., Strydom, A., Richards, M., Hardy, R., Bernal, J., & Wadsworth, M. (2005). Social outcomes in adulthood of children with intellectual impairment: Evidence from a birth cohort. *Journal of Intellectual Disability Research, 49*, 171-182.

Hettleman, K. R. (2003). *The invisible dyslexics: How public school systems in Baltimore and elsewhere discriminate against poor children in the diagnosis and treatment of early reading difficulties.* Baltimore, MD: The Abell Foundation.

Hintze, J. M, Ryan, A. L., & Stoner, G. (2003). Concurrent validity and diagnostic accuracy of the dynamic indicators of basic early literacy skills and the comprehensive test of phonological processing. *Technical Report.* University of Massachusetts Amherst. Amherst, MA.

Hitchcock, C. H., Prater, M. A., & Dowrick, P. W. (2004). Reading comprehension and fluency: Examining the effects of tutoring and video self-modeling on first-grade students with reading difficulties. *Learning Disability Quarterly, 27*, 89-102.

Hoover, J. J. & Patton, J. R. (2008). The role of special educators in a multitiered instructional system. *Intervention in School and Clinic, 43*(4), 195-202.

Hoskyn, M., & Swanson, H. L. (2000). Cognitive processing of low achievers and children with reading disabilities: A selective meta-analytic review of the published literature. *School Psychology Review, 29*(1), 102-119.

Hosp, M. K, & Saenz, L. (2005). Summer Institute: *Introduction to CBM in Reading*. Retrieved from http://www.studentprogress.org/library/Training/CBMmath/BeginingReading/beginningcbmreadingpres7-1-05.pdf

Huang, L.V., Nelson, R.B., & Nelson, D. (2008). Increasing reading fluency through student-directed repeated reading and feedback. *The California School Psychologist, 13*, 33-40.

Ingesson, S. G. (2005). Stability of IQ measures in teenagers and young adults with developmental dyslexia. *Dyslexia, 12*, 81-95.

Intelligence. (2009). *Merriam-Webster Online Dictionary*. Retrieved March 22, 2009, from http://www.merriam-webster.com/dictionary/intelligence

International Observatory on Academic Achievement. Retrieved from http://www.crires-oirs.ulaval.ca/sgc/lang/en_CA/pid/5493

Jenkins, J. R., Zumeta, R., Dupree, O., & Johnson, K. (2005). Measuring gains in reading ability with passage reading fluency. *Learning Disabilities Research and Practice*, *20*, 245-253.

Johnson, E., Mellard, D. F., & Byrd, S. E. (2005). Alternative models of learning disabilities identification: Considerations and initial conclusions. *Journal of Learning Disabilities*, *38*, 569-572.

Johnston, R. S., & Morrison, M. (2007). Toward a resolution of inconsistencies in the phonological deficit theory of reading disorders: Phonological reading difficulties are more severe in high-IQ poor readers. *Journal of Learning Disabilities, 40*(1), 66-79.

Joshi, R. M. (2003). Misconceptions about the assessment and diagnosis of reading disability. *Reading Psychology, 24*, 247-266.

Kavale, K. (2005). Identifying specific learning disability: Is responsiveness to intervention the answer? *Journal of Learning Disabilities, 38*(6), 553-562.

Kavale, K. A. (2005). Identifying specific learning disability: Is responsiveness to intervention the answer? *Journal of Learning Disabilities, 38*, 553-562.

Kim, K. H. (2005). Can only intelligent people be creative? *The Journal of Secondary Gifted Education, 2*(3), 57-66.

Knezek, G. & Christensen, R. (2007). Effect of technology-based programs on first and second-grade reading achievement. *Computers in the Schools, 24,* 23-41.

Lassiter, K. S., Matthews, T. D., & Feeback, G. (2007). An examination of the CTONI utilizing GC-GF theory: A comparison of the CTONI and WJ-III. *Psychology in the Schools, 44*(6), 567-577.

LDInfo Publishing. (2006). Retrieved from http://www.ldinfo.com/sd.htm

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "'general intelligence,' objectively determined and measured." *Journal of Personality and Social Psychology, 86*(1), 96-111.

Machek, G. R., & Nelson, J. M. (2007). How should reading disabilities be operationalized? A survey of practicing school psychologists. *Learning Disabilities Research and Practice, 22*(2), 147-157.

Magnolia Consulting, LLC. (2008) *A Study of Pearson's My Sidewalks Program: Final Report*. Louisa, Virginia: Author.

Mather, N., & Woodcock, R. W. (2001). Examiner's Manual. Woodcock Johnson III Tests of Cognitive Abilities. Itasca, IL: Riverside Publishing.

Mastropieri, M. A., & Scruggs, T. E. (2005). Feasibility and consequences of response to intervention: Examination of the issues and scientific evidence as a model for the identification of the individuals with learning disabilities. *Journal of Learning Disabilities, 38*, 525-531.

Marston, D. (2005). Tiers of intervention in responsiveness to intervention: Prevention outcomes and learning disabilities identification patterns. *Journal of Learning Disabilities*, 38, 539-544.

McComas, J. J., Wagner, D., Chaffin, M.C., Holton, E., McDonnell, M., & Monn, E. (2009) Prescriptive analysis: Further individualization of hypothesis testing in brief experimental analysis of reading fluency. *Journal of Behavioral Education,* 18, 56-70.

Mesmer, H. E. (2010). Textual scaffolds for developing fluency in beginning readers: Accuracy and reading rate in qualitatively leveled and decodable text. *Literacy Research and Instruction, 49*, 20-39.

Musti-Rao, S., Hawkins, R., & Barkley, E.A. (2009). Effects of repeated readings on the oral reading fluency of urban fourth-grade students: Implications for practice. *Preventing School Failure*, *54*(1), 12-23.

Nash, R. (2005). Cognitive habitus and collective intelligence: Concepts for the explanation of inequality of educational opportunity. *Journal of Education Policy, 20*(1), 3-21.

Nettelbeck, T., & Wilson, C. (2005). Intelligence and IQ: What teachers should know. *Educational Psychology, 25*(6), 609-630.

O'Malley, K. J., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Swank, P. R. (2002). Growth in precursor and reading-related skills: Do low achieving and IQ-discrepant readers develop differently? *Learning Disabilities Research and Practice*, *17*, 19-34.

Palmer, K., & Kevan, A. (2007). The contribution of visual sensitivity, phonological processing, and nonverbal IQ to children's reading. *Scientific Studies of Reading, 11*(1), 33-53.

Park, S., Park, K., & Choe, H. (2005). The relationship between thinking styles and scientific giftedness in Korea. *The Journal of Secondary Gifted Education, 16,* 87-97.

Pearson. (2006, April). Scott Foresman Introduces *My Sidewalks on Scott Foresman Reading Street.* Retrieved from http://www.pearsoned.com/pr_2006/042706a.htm

Restori, A. F., Gresham, F. M., & Cook, C. R. (2008). "Old habits die hard:" Past and current issues pertaining to response-to-intervention. *The California School Psychologist, 13*, 67-78.

Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality, 21*, 667-706.

Runco, M. A., Dow, G., & Smith, W. R. (2006). Information, experience, and divergent thinking: An empirical test. *Creativity Research Journal, 18*(3), 269-277.

Semrud-Clikeman, M. (2005). Neuropsychological aspects for evaluating learning disabilities. *Journal of Learning Disabilities*, *38*, 563-568.

Stage, S. A., Abbott, R. D., Jenkins, J. R., & Berninger, V. W. (2003). Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ-word reading discrepancy: Failure to validate discrepancy method. *Journal of Learning Disabilities*, *36*, 24-33.

Stanovich, K. E. (2005). The future of a mistake: Will discrepancy measurement continue to make the learning disabilities field a pseudoscience? *Learning Disability Quarterly, 28*, 103-106.

Stecker, P. M., Fuchs, D., & Fuchs, L. S. (2008). Progress monitoring as essential practice within response to intervention. *Rural Special Education Quarterly, 27*(4), 10-17.

Sternberg, R. J. (2006). Successful intelligence: Toward a broader model for teaching and accountability. *Edge: The Latest Information for the Education Practitioner, 1(5)*, 3-18.

Sternberg, R. J., & Zhang, L. (2005). Styles of thinking as a basis of differentiated instruction. *Theory into Practice*, *44*, 245-253.

Svensson, I., & Jacobson, C. (2006). How persistent are phonological difficulties? A longitudinal study of reading retarded children. *Dyslexia, 12*(1), 3-20.

Swanson, H. L., & Howard, C. B. (2005). Children with reading disabilities: Does dynamic assessment help in the classification? *Learning Disability Quarterly, 28*, 17-34.

Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities, 39(*3) 252-269.

Ukrainetz, T. A. (2006). The implications of RTI and EBP for SLPs: Commentary on L. M. Justice. *Language, Speech, and Hearing Services in Schools, 37*, 298-303.

U.S. Department of Education. (2008). IDEA Regulations. Retrieved from http://idea.ed.gov/explore/view/p/%2Croot%2Cdynamic%2CTopicalBrief%2C23%2C

U.S. Department of Education. (2006). Office of Special Education and Rehabilitation Services. Retrieved January from http://www.ed.gov/about/offices/list/osers/osep/index.html

Vadasy, P. F., Sanders, E. A., & Tudor, S. (2007). Effectiveness of paraeducator-supplemented individual instruction: Beyond basic decoding skills. *Journal of Learning Disabilities, 40(6)*, 508-525.

White, J. (2002). Puritan intelligence: The ideological background to IQ. O*xford Review of Education, 31*(3), 423-442.

Zang, L. (2002). Thinking styles: Their relationships with modes of thinking and academic performance. *Educational Psychology, 22*, 331-348.

Appendix A: Permission to use DIBELS

Dear Christine,

Thank you for your interest in Sopris West. I received your request for permission to use DIBELS *(Dynamic Indicators of Basic Early Literacy Skills),* fourth grade Oral Reading Fluency measures to identify participants for your dissertation. We are happy to grant you permission to use these pages for no charge.

In exchange for using an excerpt, we ask that you use the following reference under each use throughout the materials:
Reprinted with permission from Sopris West Educational Services. DIBELS (Dynamic Indicators of Basic Early Literacy Skills), by Roland Good and Ruth Kaminski |↵| 2003.

If you have any questions/comments, please feel free to be in touch.

Thank you,

Jillian Stanko
Sopris West Educational Services
4093 Specialty Place
Longmont, CO 80504
jills@sopriswest.com

Appendix B: Permission to use Read Naturally

Christine Tucker has permission to use the Read Naturally strategy for her doctoral study.

Thomas M. Ihnot
CEO

Read Naturally, Inc.
750 S. Plaza Dr. #100
St. Paul, MN 55120
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 – fax

Appendix C: Assent Form

ASSENT FOR PARTICIPATION IN DOCTORAL RESEARCH STUDY

Hello,

My name is Mrs. Tucker, and I am doing a research project to learn about ways to help children learn to read. I am inviting you to join in my project. I am inviting you because your recent DIBELS scores tell me that you need some help learning to read more fluently. I am going to read this form to you because I want you to learn about the project before you decide if you want to be in it.

ABOUT THE STUDY:
If you choose to be part of the study, this is what will happen:

1. All of the students will come to my room one at a time to take a quick reading test called the Gray Oral Reading Test. This should take about 15 minutes.

2. After everyone has taken the test, I will split all of the volunteers into two groups by putting everyone's name in a box and pulling names out without looking. I will do it this way to make the groups and the project fair.

3. For 10 weeks, one of the groups will come to the computer lab Tuesday through Thursday for 30 minutes to use the Read Naturally reading program on the computer. I will show you how it works, and once you understand how it works you will be able to do it yourself most of the time. There are a few parts of the program that I will need to help you with, but I will be there in the computer lab with you the whole time in case you get stuck or have any questions.

4. The other group of students will stay in the classroom and follow the regular schedule. Even though these students will not be using the program, they will still be very helpful to me. I will be able to use their testing information to decide if this computer program works.

5. After the 10 weeks, all of the students will come back to my room one at a time to take the quick reading test again, except this test will be a little bit different from the first test.

6. Then the project is over. I will be able to look at all of the scores on the reading tests to figure out if the reading program really helps kids read more fluently. If it does, all of the students in the school who need some help with reading will be able to use this reading program to help them.

IT'S YOUR CHOICE:
You don't have to be in this project if you don't want to. You won't get into trouble with anyone at the [redacted] School if you say no. If you decide now that you want to join the project, you can still change your mind later. If you want to skip some parts of the project, just tell me.

If you decide to be in this project and you are in the computer lab group, you might miss some lessons or activities in your classroom. I will talk to your teacher to make sure she sets aside anything you missed during that time. But this project might help others by helping them figure out the best ways to help kids learn to read.

There is no compensation for being in this study.

PRIVACY:
Everything you tell me during this project will be kept private. That means that no one else will know your name, what answers you gave, or how you did on the testing and on the reading program. The only time I have to tell someone is if I learn about something that could hurt you or someone else.

ASKING QUESTIONS:
You can ask me any questions you want now. If you think of a question later, you or your parents can reach me at [redacted]. If you or your parents would like to ask my university a question, you can call Dr. Leilani Endicott. Her phone number is 1-800-925-3368, then dial 1210.

Thank you very much for taking the time to learn about my study! I will give you a copy of this form. Please sign your name below if you want to join this project.

Name of Child _____

Child's Signature _____

Date _____

Researcher's Signature _____

Appendix D: Consent Form

CONSENT FOR PARTICIPATION IN DOCTORAL RESEARCH STUDY

Dear Families,

Christine Tucker, a doctoral student at Walden University and speech-language pathologist at the [redacted] School, is inviting your child to participate in a 10 week reading fluency study. This study will help teachers and administrators determine the most effective and appropriate reading interventions that enable students to improve reading fluency skills. Your child has been chosen to participate in this study based on his/her most recent Dynamic Inventory of Basic Early Literacy Skills (DIBELS) Oral Reading Fluency scores, which indicates that he/she needs improvement in the area of reading fluency.

Parent/Guardian and Child Rights

Your participation in this study is voluntary. This means that everyone will respect your decision of whether or not you want your child to be in the study. No one at the [redacted] School will treat you or your child differently if you decide not to let your child be in the study. If you decide to give your child permission to join the study now, you or your child can still change your mind later. Any children who feel stressed during the study may stop at any time. There is no compensation for being in this study. Any information your child provides in this study will be kept confidential. Mrs. Tucker will not use your child's information for any purposes outside of this research project. Also, Mrs. Tucker will not include your child's name or anything else that could identify your child in any reports of the study.

Study Description

Once participants have been established based on DIBELS Oral Reading Fluency scores, participants will be randomly divided into two groups. All participants in this study will be administered a brief assessments called the Gray Oral Reading Test: Fourth Edition (GORT:4) to assess his/her reading fluency skills, which should take no longer than 20 minutes. Following the assessment, one of the groups of students will participate in an individualized reading intervention outside of the classroom 4 days/week for 30 minutes over a 10 week period using the Read Naturally reading intervention program. This program is a research based reading intervention that many teachers throughout the country currently use with their students. The other group of students will follow the regular education reading curriculum with no changes in their schedules. Following the 10 week period, your child will be re-administered the above mentioned assessment, and pre/posttest scores will be compared and analyzed to determine your child's reading progress. Having your child participate in this study poses the minimal risk of missing some instruction in his/her classroom. However, the intervention will take place first thing in the morning, therefore minimal class work will be missed. Mrs. Tucker will make every effort to make sure students are able to make up what is missed. Potential benefits include: individualized attention, instruction, and progress monitoring in reading, which may help improve your child's reading skills.

Contact Information

You may ask any questions you have now. Or if you have questions later, you may contact Mrs. Tucker directly at [redacted]. If you want to talk privately about your rights as a participant,

you can call Dr. Leilani Endicott. She is the Walden University representative who can discuss this with you. Her phone number is 1-800-925-3368, extension 1210. Walden University's approval number for this study is 07-06-09-0345944 and it expires on July 5th, 2010. The researcher will give you a copy of this form to keep.

If you would like to grant permission for your child to participate in this study, please fill out the bottom portion below and return it as soon as possible.

Thank you very much for your time and consideration in this matter.

I have read the above information and I feel I understand the study well enough to make a decision about my child's involvement. By signing below, I am agreeing to the terms described above, and **allowing** my child to participate in this study.

Printed Name of Participant _____

Date of consent _____

Parent's Signature _____

Researcher's Signature _____

RESPONSE TO INTERVENTION: INCREASING FLUENCY, RATE, AND ACCURACY FOR STUDENTS AT RISK FOR READING FAILURE

Christine Tucker
Walden University

Don Jones
Texas A&M University - Kingsville

ABSTRACT

In recent years, with the advent of NCLB, the U S Department of Education established new guidelines for the identification of students with learning disabilities. In an effort to meet these new guidelines, many districts have implemented a Response to Intervention (RTI) model. This model has been implemented in conjunction with, or in place of the previously common IQ-achievement discrepancy model that emphasized IQ in the identification process. This quasi-experimental study examined the effectiveness of RTI by measurement of the impact of specialized instruction on the reading fluency of identified fourth grade children at risk of failure. The results of the study indicated that student achievement significantly increases when RTI is effectively implemented.

Introduction

The IQ-achievement discrepancy model was developed in 1977 as part of a federally mandated definition of learning disability and is based on the theory that "achievement predicts intelligence, intelligence is a static characteristic, and intelligence serves as a measure of learning capacity" (O'Malley et al., 2002, p. 32). In other words, this model asserts that any child who has been determined to have a low IQ with commensurate achievement is performing up to his or her learning potential, and therefore is not capable of achieving greater success through learning interventions.

Consequently, using this model has caused many students who are at risk for failure to be denied services to increase achievement (Restori, Gresham, & Cook, 2008).

The IQ-achievement discrepancy model has been used to identify learning disabilities (LD) throughout the United States since that time and in recent years has become a controversial issue in education (Machek & Nelson, 2007). Many researchers have argued that IQ is the most critical aspect of identifying children with LD (Johnson, Mellard, & Byrd, 2005; Kavale, 2005) while others have suggested that IQ is just a small piece of the puzzle, and more focus should be placed on dynamic assessments and responsiveness to intervention (Cooter & Cooter, 2004; Fletcher, Francis, Morris, & Lyon, 2005; Hettleman, 2003; National Joint Committee on Learning Disabilities, 2005; Fusch et al. 2003).

Because of this conflict, many researchers have investigated different models of identifying LD, which spurred the U.S. Department of Education (USDOE) to create new guidelines in the identification process. In 2004, President George W. Bush signed a reauthorization of the Individuals with Disabilities Education Act (IDEA) that stated:

> A state must adopt, consistent with 34 CFR 300.309, criteria for determining whether a child has a specific learning disability as defined in 34 CFR 300.8(c)(10). In addition, the criteria adopted by the State: 1) Must not require the use of a severe discrepancy between intellectual ability and achievement for determining whether a child has a specific learning disability, as defined in 34 CFR 300.8(c)(10); 2) Must permit the use of a process based on the child's response to scientific, research-based intervention; and 3) May permit the use of other alternative research-based procedures for determining whether a child has a specific learning disability, as defined in 34 CFR 300.8(c)(10). (U.S. Department of Education, 2008, ¶ 2)

After this reauthorization, school systems across the country began implementing a response to intervention (RTI) model to use in lieu of or in conjunction with the IQ-achievement discrepancy model. However, because this newer method of identification is still optional, many school systems continue to use the discrepancy model (Restori, Gresham, & Cook, 2008).

The purpose of this study was to investigate the effect of RTI by determining the impact of specialized instruction on reading fluency skills with students in Grade 4 who were identified as at risk for reading failure. These students have been struggling to achieve academic success throughout their educational careers and some have not been given the opportunity to learn through nontraditional means. Many reading interventions have been used with students with LD to increase achievement; however, students who have been identified as slow learners or non-learning disabled based on the IQ-achievement discrepancy model have not been able to participate in such interventions (Cooter & Cooter, 2004; Kavale, 2005; Kavale, Holdnack, & Mostert, 2005). This study attempted to determine if a group of students who are at risk for reading failure was able to increase reading achievement using an evidence-based reading intervention. Because reading rate and accuracy are so closely related to reading fluency, these areas were also explored in this study.

Theoretical Framework

Part of the theoretical framework for this study is the intelligence theory, which asserts that intelligence, or IQ, predicts an individual's ability to learn. This theory, developed by Spearman (1904), is also known as the theory of general intelligence or Spearman's g and is based on the assumption that intelligence is an unchanging characteristic that determines learning capacity (Lubinski, 2004; O'Malley et al., 2002). The theory of general intelligence has been the basis for how the special education system has operated for more than 30 years (Fletcher, Francis, Morris, & Lyon, 2005; Francis

et al., 2005; Fuchs et al., 2003). Because the IQ-achievement discrepancy model is based on this theory, it served as part of the theoretical framework for this study

Additionally, the theory of dual discrepancies presented by Fuchs, Fuchs, and Speece (2002) also constituted the theoretical framework for this study. This theory states that "when a low-performing child fails to manifest growth in a situation where others are thriving, alternative instructional methods must be tested to address the apparent mismatch between the student's learning requirements and those represented in the conventional instructional program" (p. 35). This theory, essentially, is the foundation for the RTI model, which was implemented in this study. Both the intelligence theory and the dual discrepancies theory played vital roles in this study.

Methodology

This quasi-experimental quantitative study, using a pretest and posttest control group design, investigated the impact of a research based reading intervention on reading fluency, rate, and accuracy skills among fourth-grade students who were at risk for reading failure, as evidenced by the students' DIBELS Oral Reading Fluency scores. This study attempted to answer the following research questions:

1. Is there a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group?
2. Is there a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group?
3. Is there a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group?

Research Design

A quasi-experimental, quantitative, pretest and posttest control group design was chosen due to the nature of the research questions. According to Creswell (2003), "If the problem is identifying factors that influence an outcome, the utility of an intervention, or understanding the best predictors of outcomes, then a quantitative approach is best" (p. 21). Creswell also indicated that "In quasi-experiments, the investigator uses control and experimental groups" (p. 167). The pre and posttests enabled the researcher to determine each group's progress in reading fluency, rate, and accuracy to draw conclusions about the effect of the intervention.

Participants and Setting

This study took place in an elementary school in a large urban district in the northeastern region of the United States. This study used a convenience sample to identify potential participants. A convenience sample was used because the researcher needed to identify participants based on current data, more specifically, the most current fourth-grade DIBELS Oral Reading Fluency scores. The sample consisted of fourth-grade students who placed in the at risk category based on their most recent DIBELS Oral Reading Fluency scores. Fourth-graders were chosen as the targeted population because research shows that many times, reading disabilities often do not present themselves and are not diagnosed until fourth-grade (Lipka, Lesaux, & Siegel, 2006). The National Center for Education Statistics (NCES) found that 40% of fourth graders did not meet the basic requirements set forth by the National Assessment of Educational Progress in Reading (NCES, 2006).

Twenty participants were chosen from the sample based on the following eligibility criteria: (a) students must have been in the fourth-grade, (b) students must have been in the at-risk category in reading based on DIBELS Oral Reading Fluency scores, (c) students

must not have been identified as having a specific learning disability in the area of reading, and (d) students must not have been receiving supplemental reading instruction outside of the general education classroom. The decision to use 20 participants was derived based on the relatively small sample size according to current DIBELS Oral Reading Fluency data. The relatively small number of participants allowed the researcher to synthesize and analyze the data for each student in depth to draw sound conclusions about their response to the intervention provided.

Participants' Rights

In order to protect participants' rights, each participant in this study and their legal guardians were provided with a detailed written description of the study including the purpose of the study, details regarding the reading intervention program, potential benefits, and the minimal potential risks. Legal guardians were provided with a letter of consent that they were required to sign if they chose to have their child participate in the study. The form contained contact information regarding where and when the researcher could be reached to field any questions the guardians or the participants may have had regarding the study. Legal guardians and their children's participation were voluntary, and they could have removed their child from the study at any time. Personal information was not used in this study; each student's information was coded to maintain confidentiality, and students' names were changed as well. All student data were stored on the researcher's computer, which required a password to access to protect the participants.

Limitations and Boundaries

This study had limitations that make it difficult to generalize the results to the general population. One important limitation was the relatively small sample size used in this study. The goal of any

research study is to generalize the results from a selected sample back to the entire population (Gravetter & Wallnau, 2005). However, because this study used a sample of only 20 students, additional research corroborating the results would be beneficial in generalizing the results to the population.

Another limitation of this study was that it focused solely on reading fluency and its related components. In order to truly determine if RTI is effective, other areas of learning should be investigated. Moreover, this study was limited to fourth grade students. If this study had used participants from other grade levels the results may have been different. To validate the effects of RTI, interventions should be implemented and studied with students from various grade levels. Also, this study only used one reading intervention. The results may have been different if a variety of reading interventions were used. Consequently, it is necessary to investigate other reading programs to determine if RTI is effective using various interventions.

In addition to the limitations posed by the design of the study, the researcher also could not account for the behavior of the participants in this study. The participants had varying levels of motivation, and therefore they each demonstrated varying amounts of effort and dedication to improving their reading fluency skills. Some students were highly motivated and were extremely focused when participating in the intervention while other students were disinterested and simply went through the motions. Also, it is unknown how much time students in the study practiced reading outside of the allotted time, which may have impacted some of the student's posttest scores.

Data Collection

Following random assignment to either the control group or the quasi-experimental group, each participant was tested (pretest) using form A of the GORT-4 to assess reading fluency, rate, and accuracy. Following the pretest, the students in the control group received

reading instruction using the general education curriculum within the regular education classrooms, and the experimental group received supplementary reading instruction, in addition to the general education reading instruction, using the Read Naturally reading intervention program outside of the regular education setting 4 days per week for 30 minutes across 10 weeks. Following the 10-week period, all participants were retested using form B of the GORT-4 to assess reading fluency (posttest).

Data Analysis

Following the posttest, an independent measures *t* test was used to evaluate the mean difference between the experimental group's growth in the area of fluency, rate, and accuracy based on pretest and posttest scores and the control group's growth in the area of fluency, rate, and accuracy based on pretest and posttest scores. Because the independent measures *t* test only measures the treatment effect, the percentage of variance accounted for by the treatment was also calculated. According to Gravetter and Wallnau (2005), this measure examines whether the treatment causes the scores to vary. In addition, Hartley's F-max test was also used to satisfy the homogeneity of variances assumption, which indicates that the two populations that are being compared have the same variances. Gravetter and Wallnau (2005) indicated that this is an important calculation because if the assumption is not met, the data cannot be interpreted meaningfully from an independent measures experiment.

Fluency

The following hypotheses were tested using an independent measures *t* test.

Ho_1: There is no significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group..

Ha_1: There is a significant difference in the reading fluency growth based on pretest and posttest scores between the control group and the quasi-experimental group.

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.29 was obtained, which falls within the critical region. Because the *t* score value fell within the critical region, the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest scores of the experimental group is significantly higher than the growth based on pretest and posttest scores of the control group. Figure 1 shows the fluency pretest and posttest scores for each student in the control group, and Figure 2 shows the fluency growth based on pretest and posttest scores for each student in the control group.

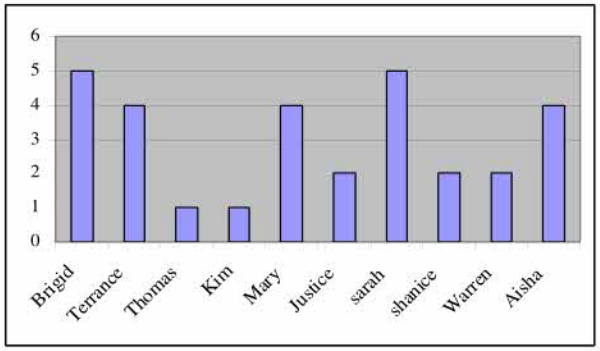


Figure 1: *Experimental group fluency growth*

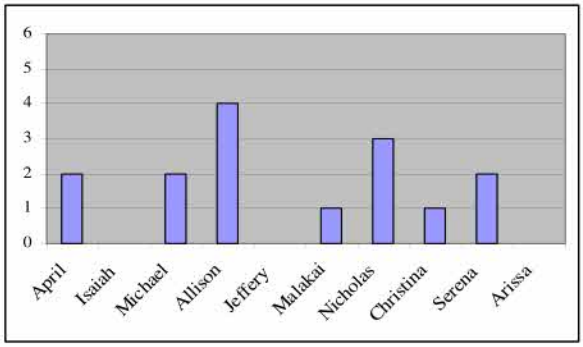


Figure 2: *Control group fluency growth*

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 3.00 with a SD = 1.56 between pretest and posttest scores. The students in the control group had a M = 1.50 with a SD = 1.35 between pretest and posttest scores. This difference was significant, $t(18) = 2.29$, $p < .05$, $r^2 = .23$, which indicates that the 10-week reading intervention using the Read Naturally program had a medium effect on the students' reading fluency scores.

The F-max (1.55) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Rate

The following hypotheses were tested using an independent measures *t* test.

> Ho_1: There is no significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Ha_1: There is a significant difference in the reading rate growth based on pretest and posttest scores between the control group and the quasi-experimental group.

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.51 was obtained, which fell within the critical region. Because the *t* score value fell within the critical region, the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest rate scores of the experimental group is significantly higher than the growth based on pretest and posttest rate scores of the control group. Figure 3 shows the rate pretest and posttest scores for each student in the control group, and Figure 4 shows the rate growth based on pretest and posttest scores for each student in the control group.

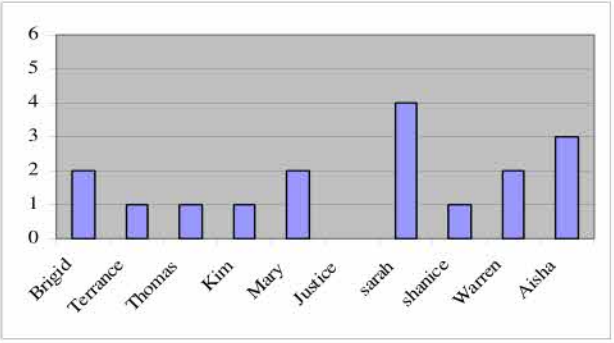


Figure 3: Experimental group rate growth

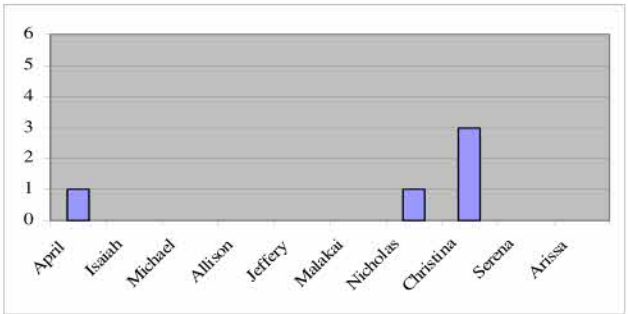


Figure 4: Control group rate growth

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 1.70 with a SD = 1.16 between pretest and posttest scores. The students in the control group had a M = 0.50 with a SD = 0.97 between pretest and posttest scores. This difference was significant, $t(18) = 2.51$, $p < .05$, $r^2 = .26$, which indicates that the 10-week reading intervention using the Read Naturally program had a large effect on students' reading rate scores.

The F-max (1.43) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Accuracy

The following hypotheses were tested using an independent measures *t* test.

> Ho_1: There is no significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Ha_1: There is a significant difference in the reading accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group..

In a one-tailed independent measures *t* test with an alpha level of .05 and a *df* of 18, any *t* score value greater than +1.734 is considered to be within the critical region. A *t* score value for the sample mean of 2.89 was obtained, which fell within the critical region. Because the *t* score value fell within the critical region the null hypothesis was rejected. The conclusion is that the growth based on pretest and posttest rate scores of the experimental group is significantly higher than the growth based on pretest and posttest rate scores of the control group. Figure 5 shows the accuracy pretest and posttest scores for each student in the control group, and Figure 6 shows the accuracy growth based on pretest and posttest scores for each student in the control group.

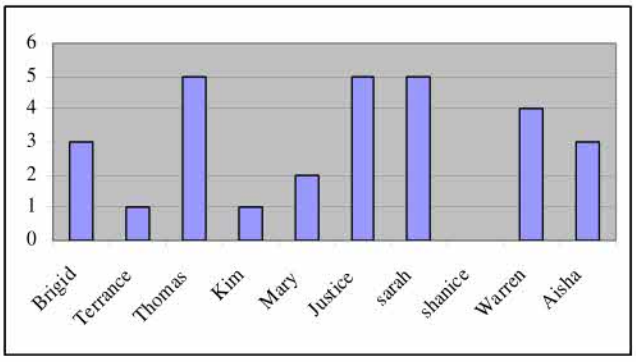


Figure 5: Experimental group accuracy growth

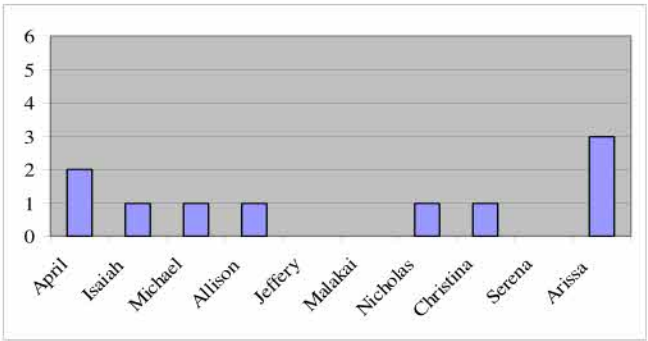


Figure 6: Control group accuracy growth

The sample size of the pretest and posttest was $n = 10$ for each group. The students in the experimental group had a M = 2.90 with a SD = 1.85 between pretest and posttest scores. The students in the control group had a M = 1.00 with a SD = 0.94 between pretest and posttest scores. This information is found on Table 11. This difference was significant, $t(18) = 2.89$, $p < .05$, $r^2 = .32$, which indicates that the 10-week reading intervention using the Read Naturally program had a large effect on the students' reading rate scores.

The F-max (3.49) calculated for the individual samples fell below the critical value of 4.03. Therefore, the data suggest that the population variances are similar and the homogeneity assumption is reasonable.

Summary

According to the independent measures *t* tests that were conducted, the difference between the experimental group's growth and the control group's growth is significant in all three areas that were examined. Pretest and posttest results indicate that students in the experimental group significantly improved their fluency, rate, and accuracy following the 10-week reading intervention. Although the control group also improved in these areas, the improvement was not

considered to be significant. The null hypotheses, which stated that there is no significant difference in the reading fluency, rate, and accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group, were rejected. Therefore, the answers to the research questions are yes: There is a significant difference in the reading fluency, rate, and accuracy growth based on pretest and posttest scores between the control group and the quasi-experimental group.

Recommendations

Implementing the suggestions presented earlier regarding future research could allow the educational system to make sound resolutions on the best way to increase student achievement and identify students with special needs. It will also provide educational systems with the impetus to adopt this method. Administrators in special education departments who determine qualifications for specialized instruction should pay careful attention to research in this area to aid in making decisions regarding the use of RTI. Bergstrom (2008) argued that successfully adopting and implementing RTI goes far beyond progress monitoring and scientifically based interventions; it requires a comprehensive school wide system reform. According to Berkeley, Bender, Peaster, and Saunders (2009), only 15 states have fully developed and are utilizing RTI models. Therefore, the USDOE should look at the results of this study as well as future research to aid in determining the most effective method of identifying learning disabilities, and ultimately implementing new special education standards for all school systems.

The data from this study support RTI as an effective way to improve student achievement. Educators who have students who are at risk for reading failure should also pay attention to the results of this study as well as future research to assist in determining methods of instruction and development of lesson plans and curriculum. The results from this study could be shared with teachers and other faculty

members through professional development to provide educators with a concrete example of how valuable this approach can be with struggling students.

In order for this method to be implemented effectively, educators must develop a deep understanding of and belief in this process. Furthermore, providing teachers and other faculty with the appropriate training will allow them to build confidence in their ability to use RTI with their students. Bergstrom (2008) asserted that in order for RTI to be successful, one of the most important pieces in the reform is professional development geared towards educating teachers about the process. Without the appropriate professional development, the execution of such a program would be impossible.

Conclusion

As stated earlier, the USDOE has created new guidelines for identifying learning disabilities that allow school systems to adopt the RTI method in lieu of or to use in conjunction with the IQ-achievement discrepancy model. Although RTI has been implemented in some school systems throughout the United States, Carney and Stifel (2008) pointed out that "While RTI has come to schools through federal legislation, it has been left to educators and researchers to interpret and investigate the best means of operationalizing this intent to ensure that students difficulties do not stem from instructional deficiencies" (p. 61). In other words, the vague wording of the changes in IDEA leave it open to interpretation. It is up to individual school systems to make decisions on how to structure RTI and ensure students who are at risk are provided with opportunities to become successful. This study, along with future research, will allow the Department of Education, school administrators, and other educational leaders to evaluate the RTI process and establish a protocol for its use in schools across the country to improve student achievement, ultimately effecting social change.

Without relevant research on RTI, many school systems would continue to use the IQ-achievement discrepancy model and perpetuate the ongoing problem of students who do not qualify for special education services falling through the cracks. These students may not be given the chance to increase their achievement skills, even though they struggle with the curriculum simply because they do not qualify to receive specialized instruction (Ukrainetz, 2006). Implementing RTI in schools across the country will allow all students the opportunity to participate in interventions geared towards their specific learning needs. With this targeted instruction, students may be able to improve very precise areas of deficit to advance their overall educational performance, which in turn will afford them greater opportunities in the future.

REFERENCES

Bergstrom, M. K. (2008). Professional development in response to intervention: Implementation of a model in a rural region. *Rural Special Education Quarterly, 27*(4), 27-36.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities, 42*(1), 85-95.

Carney, K. J., & Stiefel, G. S. (2008). Long-term results of problem solving approaches to response to intervention: Discussions and implications. *Learning Disabilities: A Contemporary Journal, 6*(2), 61-75.

Cooter, K. S., & Cooter, R. B. (2004). One size does not fit all: Slow learners in the reading classroom. *The Reading Teacher, 57*(7), 680-684.

Creswell, J. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.

Gravetter, F.J. & Wallnau, L.B. (2005). *Essentials of statistics for the behavioral sciences* (5th ed.). New York, NY: Thomson Publishing.

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34,* 506-522.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities*, *38*, 98-108.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disabilities Quarterly, 25*(1), 33-45.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness to intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18*, 157-171.

Hettleman, K. R. (2003). *The invisible dyslexics: How public school systems in Baltimore and elsewhere discriminate against poor children in the diagnosis and treatment of early reading difficulties.* Baltimore, MD: The Abell Foundation.

Johnson, E., Mellard, D. F., & Byrd, S. E. (2005). Alternative models of learning disabilities identification: Considerations and initial conclusions. *Journal of Learning Disabilities*, *38*, 569-572.

Kavale, K. A. (2005). Identifying specific learning disability: Is responsiveness to intervention the answer? *Journal of Learning Disabilities, 38*(6), 553-562.

Kavale, K.A., Holdnack, J.A., & Mostert, M.P. (2005). Responsiveness to intervention and the identification of specific learning disability: A critique and alternative proposal. *Learning Disability Quarterly, 28*(1), 2-16.

Lipka, O. Lesaux, N. K., Siegel, & L. S. (2006). Retrospective analyses of the reading development of grade 4 students with reading disabilities: Risk status and profiles over 5 years. *Journal of Learning Disabilities, 39*(5), 364-378.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "'general intelligence,' objectively determined and measured." *Journal of Personality and Social Psychology, 86*(1), 96-111.

Machek, G. R., & Nelson, J. M. (2007). How should reading disabilities be operationalized? A survey of practicing school psychologists. *Learning Disabilities Research and Practice, 22*(2), 147-157.

National Joint Committee on Learning Disabilities (2005). *Response to Intervention (RTI): A primer for parents.* Retrieved from http://www.ldonline.org/article/15857

O'Malley, K. J., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Swank, P. R. (2002). Growth in precursor and reading-related skills: Do low achieving and IQ-discrepant readers develop differently? *Learning Disabilities Research and Practice, 17*, 19-34.

Restori, A. F., Gresham, F. M., & Cook, C. R. (2008). "Old habits die hard:" Past and current issues pertaining to response-to-intervention. *The California School Psychologist, 13*, 67-78.

Ukrainetz, T. A. (2006). The implications of RTI and EBP for SLPs: Commentary on L. M. Justice. *Language, Speech, and Hearing Services in Schools, 37*, 298-303.

U.S. Department of Education. (2008). IDEA Regulations. Retrieved August 19, 2008 from http://idea.ed.gov/explore/view/p/%2Croot%2Cdynamic%2CTopicalBrief%2C23%2C

WhatWorks

From:WhatWorks
Sent:15 Jun 2012 13:54:15 +0000
To:Info@readnaturally.com
Subject:RE: Beginning Reading Reevaluation (WWC 3702)
Dear Mr. Ihnot,

Thank you for your email. We have forwarded the Tucker Study to the review team.

The WWC Quality Review Team (QRT) is finalizing the written response to your recent inquiries. After receiving the QRT response, you may ask any additional questions in writing via the Help Desk.

Due to the need to maintain a written record of interactions and key decisions, the WWC requests that all questions are submitted in writing.

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: info@readnaturally.com [mailto:Info@readnaturally.com]
Sent: Thursday, June 14, 2012 12:10 PM
To: What Works
Subject: Beginning Reading Reevaluation

Beginning Reading Reevaluation,

I have attached the Tucker study of Read Naturally to be evaluated by What Works Clearinghouse (WWC).

Also, since there are serious issues with the four studies WWC has posted as reviews of Read Naturally, I think I should have a chance to meet with a member of the quality review team.

http://www.readnaturally.com/approach/wwc-summary.htm

Please email me a time and place for me to meet with the quality review team to discuss the four studies that are providing inaccurate information to educators.

Thank you,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190

St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

February 26, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2010013

Dear Mr. Ihnot:

Thank you for your inquiries concerning the What Works Clearinghouse (WWC) reviews of four studies of Read Naturally®—Kemp (2006), Denton et al. (2004), Hancock (2002), and Chenault et al. (2006)—and for your email regarding my response to Dr. Kevin Feldman's inquiry about the WWC reviews of studies of Read Naturally®. In response to your concerns, we conducted an independent quality review. The WWC quality review team responds to concerns raised by study authors, curriculum developers, or other relevant parties about WWC reviews published on our website. When a quality review is conducted, a researcher who was not involved in the initial reviews undertakes an independent assessment of the studies in question. The researcher also investigates the procedures used and decisions made during the original reviews of the studies. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the WWC.

The quality review considered the concerns from your emails regarding study inclusion, implementation, student sample, and outcomes.

Study Inclusion

The WWC followed protocol in choosing to review these four studies. The WWC screens studies based on a number of factors, including relevancy and methodology criteria (as described in the WWC Procedures and Standards Handbook, pages 8–10, available at http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19). The WWC does not screen based on whether the author(s) explicitly intended the study as an evaluation of an intervention, but rather whether the study presents a primary analysis of the effect of an intervention. This screening process allows for the review of a wide range of relevant and methodologically sound studies.

Thank you for bringing to my attention the studies by Mesa (2004) and Tucker (2010). The Tucker (2010) study will be included in our literature search for updating the Intervention Report

for Read Naturally®, reviewed under the Beginning Reading Evidence Review Protocol. In addition, as you note, the Mesa (2004) study was included in the original review of Read Naturally®, but will be reviewed again in an updated review.

Implementation

The WWC followed protocol in the manner in which the four studies are described in WWC publications. The WWC Handbook says, "The WWC makes no adjustments or corrections for variations in implementation of the intervention; however, if a study meets standards and is included in an intervention report, descriptions of implementation are provided in the report appendices to provide context for the findings" (page 16). This approach is appropriate because there is no standard metric for fidelity to intervention design. Thus, the WWC includes studies with variation in fidelity and does not evaluate implementation fidelity. The quality review team verified that variations in implementation that are noted in the four studies and that may affect the interpretation of findings were properly included in the WWC publications.

Regarding the Kemp (2006) study, your emails raised a concern about implementation of the Read Naturally® program. The quality review noted that the study does not provide any indication of deviation from program design. Specifically, the study states, "… it could be concluded that all teachers implemented the Read Naturally® program as prescribed" (page 40). Based on this information, the quality review concluded that there was no evidence that deviations in implementation should have been noted in the WWC Intervention Report. Regarding your specific concern about the teacher-to-student ratio, the study does not note any deviation from the Read Naturally® program in this area and does not provide any information on the ratio.

You also expressed concern over deviations in the implementation of Read Naturally® in the Denton et al. (2004) study. The quality review found that the inclusion of additional activities was noted in the study and reported in the WWC Intervention Report (see footnote 7 on page 3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). The other potential deviations that you noted (accountability to pass criteria, use of audio tapes, and the number of sessions per week) were not identified in the study and are therefore not noted in the WWC publication.

Regarding your concerns that the Read Naturally® strategy was not implemented fully for the Hancock (2002) study, the quality review found that WWC procedures were followed in the WWC Intervention Report. The WWC contacted Hancock regarding the implementation of Read Naturally® in the 2002 study. Her response indicated that the study excluded Read Naturally®'s pre-reading vocabulary instruction component and the placement system to individualize instruction. The WWC Intervention Report was revised to note these deviations in implementation (see footnote 4 on page 2, available at http://ies.ed.gov/ncee/wwc/InterventionReport.aspx?sid=407). Other potential concerns with implementation described in your email were not reported by the author and were not included in the revised WWC Intervention Report.

The quality review also investigated concerns that you raised in prior emails. The email from June 2012 regarding Dr. Feldman's questions was addressed in my letter to Dr. Feldman; I will not

repeat the quality review findings here. Your email from May 2012 included some text that was reported to be a quote from Hancock's email to the WWC, but the quoted statements were not part of Hancock's email text. The misreported quote is also found on the Read Naturally® website at http://www.readnaturally.com/approach/hancock.htm.

Regarding the Chenault et al. (2006) study, you expressed concern with the implementation of Read Naturally® related to the selection of students and the limited number of sessions. The quality review found that the WWC Intervention Report appropriately notes both of these factors (see pages 2–3, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=409).

Student Sample

You also raised a concern that the Kemp (2006) study included a large share of students who were above the 50th percentile in reading fluency who were not "at-risk" readers. The quality review verified that the WWC Intervention Report accurately states that the study population for the results reported by the WWC was third-grade English language learner (ELL) students (page 2, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408). Furthermore, the quality review noted that Read Naturally® program materials state that the program "… products help students of all ages and abilities—whether special education, ELL, Title I, or mainstream—to improve their reading skills, gain confidence, and transform their attitudes toward school and learning" (available at http://www.readnaturally.com/products/default.htm).

Finally, regarding the differences in pre-test performance among native-speaking students in the treatment and control groups, the quality review found that these concerns are not relevant to this WWC publication, which reports on findings only for English language learner students. Furthermore, the quality review concluded that the WWC appropriately restricted the reported findings to impacts assessed relative to the comparison group, and thus did not include achievement gains in oral reading fluency.

Outcomes

Regarding the Denton et al. (2004) study, you expressed concern that the study did not assess fluency. The quality review noted that the WWC Intervention Report clearly states that the study assessed reading achievement (page 4, available at http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408), an outcome of relevance to the WWC ELL topic area. The quality review noted that the Read Naturally® program materials state there are several components of the program beyond fluency that are related to reading achievement: "Our industry-leading programs develop and support the five essential components of reading, identified by the National Reading Panel: phonemic awareness, phonics, fluency, vocabulary, and comprehension" (see "The Read Naturally Strategy," available at http://www.readnaturally.com/approach/default.htm).

I hope that this letter has addressed your concerns. You requested that we meet to discuss the WWC reviews of Read Naturally®. However, WWC policy requires all correspondence in writing. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov.

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

cc: Joy Lesnick, IES

Sent:26 Feb 2013 20:09:42 +0000
To:'kramy@readnaturally.com';'info@readnaturally.com'
Subject:What Works Clearinghouse (QR2010013)
Attachments:QRT2010013.pdf
Dear Mr. Ihnot:

Attached is a response to the questions you raised in your inquiries concerning the What Works Clearinghouse reviews of studies of Read Naturally®.

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/.

School Psychology Review,
2010, Volume 39, No. 3, pp. 447–462

Curriculum-Based Measurement of Oral Reading: An Evaluation of Growth Rates and Seasonal Effects Among Students Served in General and Special Education

Theodore J. Christ
University of Minnesota

Benjamin Silberglitt
Technology and Information Educational Services

Seungsoo Yeo
Korean Educational Development Institute

Damien Cormier
University of Minnesota

Abstract. Curriculum-based measurement of oral reading (CBM-R) is often used to benchmark growth in the fall, winter, and spring. CBM-R is also used to set goals and monitor student progress between benchmarking occasions. The results of previous research establish an expectation that weekly growth on CBM-R tasks is consistently linear throughout the academic year. The patterns of CBM-R growth were examined for a large sample of students (N = 3808) from both general education and special education populations in second to sixth grades. Results support four general conclusions: (a) annual growth is more substantial within the general education population; (b) growth is more substantial in earlier elementary grades; (c) more growth occurs in the fall than the spring season (i.e., seasonal effect), especially within the early primary general education population; and (d) the seasonal effect is less pronounced within the special education population. Estimates of growth within and across seasons are presented and implications are discussed.

Curriculum-based measurement (CBM) is used to index annual student growth across the primary grades. Procedures and measurement metrics are developed for mathematics, spelling, written expression, and reading. CBM oral reading (CBM-R) rate is the most researched and well established of those available procedures (Wayman, Wallace, Wiley, Ticha, & Espin, 2007). CBM-R is used as an index of reading rate and fluency, which are identified as critical skills to target for instruction and intervention within the early stages of

Correspondence regarding this article should be addressed to Theodore J. Christ, 56 East River Road, Minneapolis, MN 55455; E-mail: tchrist@umn.edu

Copyright 2010 by the National Association of School Psychologists, ISSN 0279-6015

reading development (National Reading Panel, 2000; National Research Council, 1998). CBM-R outcome also functions as a robust indicator of overall reading development throughout the primary grades (Wayman et al., 2007).

CBM-R progress monitoring data are collected and used to evaluate instructional effects, determine when to modify the instruction, and evaluate response to intervention (Deno, 1985, 1986). Teachers who use progress monitoring data are likely to use more specific and measurable goals, rely more substantially on data to guide instruction, and modify instruction more frequently (Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, & Stecker, 1989). Progress monitoring applications of CBM-R have the potential to confer the most substantial benefit to students. Progress monitoring can be conceptualized as either interim benchmark/screening assessments when data are collected three to four times per year, or time series continuous assessments when data are collected daily or weekly.

Both interim and time series CBM-R data are plotted graphically and evaluated against goals and goal lines. Fuchs and Shinn (1989) recommended that the rate of expected growth should derive from local normative performance so that individual student achievement can be compared against the local normative performance of grade-level peers. A goal line is established by connecting the CBM-R data point for the observed level of performance from Week 1 (initial performance) to the expected/typical level of performance at Week 36 (ending performance). Once graphed, the goal line provides a graphic trajectory of expected growth for the academic year. Ongoing time series data are then plotted on the graph and evaluated against the goal line. For example, if a student's initial level of CBM-R performance in September was 20 words read correctly per minute (WRCM) and in May that student was expected to read 60 WRCM, then the expected annual growth is 40 WRCM across 8 months, or 32 weeks. That translates to ~1.25 WRCM per week of growth (1.25 WRCM = 40 WRCM/32 weeks). The goal line and trajectory of expected growth is based on the assumption of a monotonic and linear trend.

Monotonic Linear Growth

Standards of expected performance within and across the year provide the foundation to establish instructional goals and goal lines. There are two influential studies that established standards of weekly CBM-R growth (Deno, Fuchs, Marston, & Shin, 2001; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993), which have application in both research and practice. Fuchs et al. (1993) analyzed a CBM-R data set comprised of 103 students in general education and 14 students in special education. Each student was assessed weekly for an entire academic year using CBM-R procedures. Although some students evidenced a negatively accelerating annual growth, the researchers concluded that annual growth is, on average, monotonic and linear. They also observed a negatively accelerating trend across grades. That is, the rate of CBM-R growth was positive in all grades, but the rate of growth declined in each successive grade. The resulting analysis yielded grade-specific estimates of standard and ambitious rates of growth—and *assumed* a pattern of monotonic linear growth. Respectively, the growth, determined by the number of WRCM, were 1.5 and 2.0 in first and second grade, 1.0 and 1.5 in third, 0.85 and 1.1 in fourth, 0.5 and 0.8 in fifth, and 0.3 and 0.63 in sixth.

Deno et al. (2001) analyzed a large data set comprised of 2675 students in general education and 324 in special education classes. The student sample was from school districts across the country. The researchers relied on the *assumption* of monotonic linear annual growth. Ordinary least squares regression was used to calculate weekly growth. The study yielded useful estimates of observed and expected growth across grades and populations. The authors estimated that beginning readers in the general education population can be expected to improve 2 WRCM/week until they achieve 30 WRCM; thereafter, students in the general education population can be expect to improve at least 1 WRCM/week. Researchers observed that students in the special education

population demonstrated substantially less growth, which approximated 0.5–0.8 WRCM/week, unless they are provided with robust evidence-based instruction, which might improve the observed rate of growth to 1.39 WRCM/week.

Prior studies provide useful estimates of typical, ambitious, and expected rates of CBM-R growth; however, those estimates depend on the assumption that growth is monotonic and linear across the academic year. That assumption should be examined more closely to determine its veracity. The prior studies also rely on analyses that neglect the inherent nesting of students within classrooms and schools, which establish potential bias in the linear and quadratic functions derived.

Improved Modeling of Annual and Seasonal Growth

Research is necessary to evaluate alternate models of annual growth and test the predominant assumption that annual growth is monotonic and linear. This study will evaluate both linear (constant) and piece wise (nonconstant) models of growth from fall to winter to spring; moreover, this study will employ the superior approach of linear mixed model (LMM) rather than that of ordinary least squares used in prior research (Deno et al., 2001; Fuchs et al., 1993).

Ordinary least squares only estimates the fixed effects, which assumes that all cases have identical parameter estimates for intercept and growth. This is a limitation when modeling CBM-R growth because it is likely to be an erroneous assumption in most cases. The results of published research provide evidence for substantial magnitudes of variance in both intercept and growth among students in a sample (Shin, Deno, & Espin, 2000; Stage, 2001). It is likely that the intercept and growth rates should be treated as random effects in many instances. LMM can be used to estimate both the fixed effects and the random effects, which are the intercepts and growth rates for individual student cases in the sample. LMM takes into account the variances within and across individuals in the sample. Finally, ordinary least squares requires that missing data are removed casewise whereas LMM can handle missing data, which is likely to yield less biased estimates of intercept and slopes while enhancing power (Duncan, Duncan, & Strycker, 2006). This approach facilitates analysis of both seasonal effects, as pursued by two previous studies (Ardoin & Christ, 2008; Graney, Missall, Martinez, & Bergstrom, 2009), and compares those effects to that of linear growth as assumed in other studies (Deno et al., 2001; Fuchs et al., 1993).

The calendar year was divided into three seasons for the purpose of this study. The seasons were defined by those periods between standard tri-annual assessment occasions: fall, winter and spring. The duration between fall and winter assessments defines the fall season, that between winter and spring defines the spring season, and that between spring and fall defines the summer season. If growth is consistent and linear throughout the calendar year, then there is no *seasonal effect*. Conversely, if growth is not consistent and linear throughout the calendar year, then there is a seasonal effect whereby there is more growth observed for some seasons than for others.

The summer season is the most distinct for the majority of students because they are not exposed to formal instruction. The rate of achievement is likely to decline in the summer months (Kim, 2004) unless formal instruction is provided through summer school programs (Stage, 2001). A seasonal effect in the summer is expected because the conditions for learning are substantially different from that which occurs in the spring and fall seasons. The effect, however, does not seem specific to those differences between summer and the academic year. Ardoin and Christ (2008) observed statistically significant differences in CBM-R growth between the fall and spring seasons. That is, among a sample of second-grade students there was more growth in the fall than in the spring season. That result replicated across three combinations of CBM-R passages sampled from the Dynamic Indicators of Basic Early Literacy Skills. A subsequent study aimed to replicate and extend those findings

with a sample of third- through fifth-grade students across a variety of CBM curriculum domains (Graney et al., 2009). However, Graney et al. (2009) concluded that there was more CBM-R growth in the spring season than in the fall season, failing to replicate the work of Ardoin and Christ. Graney et al. speculated that the failure to replicate findings across studies might relate to the specification of data collection schedules, sample characteristics, issues of instrumentation, or the distribution of instructional intensity across the academic year. Although it is difficult to make a single attribution, it should be noted that the student sample used by Graney et al. derived from a school–university partnership program that included features to promote data use and evidence-based interventions. Indeed, that program might influence the pattern of student growth; however, additional research and replications are necessary to help determine whether specific patterns for seasonal effects should be expected.

There are limitations to the analytic procedures used in the two prior studies. Both generated estimates of growth from difference scores across seasons. That is, growth was estimated as the difference between performance at Time 1 and Time 2, which was then divided by the number of intervening weeks to yield an estimate of weekly growth. Although this method is intuitive, and it was useful for the preliminary work to examine seasonal effects, the method of difference scores to estimate change has long been criticized as an unreliable method of analysis (Cronbach & Furby, 1970) and may have contributed to the inconsistent findings. The analytic procedures used in this study improve on those of prior studies that examined seasonal effects and modeled annual growth. The analytic procedures used in this study will compare the hypothesized seasonal effect directly with that of linear growth.

Purpose

The purposes of this study were to (a) estimate the rate of growth within and across seasons for general education and special education populations; (b) evaluate whether the rate of CBM-R growth is likely to be consistent (i.e., linear model) or inconsistent (i.e., piecewise model) throughout the academic year; and (c) test the hypothesis that the seasonal effect is typically characterized by more robust growth in the fall than for the spring season. That hypothesis is counter to the assumptions inherent to the linear models as applied in many previous studies and it is counter to the findings of Graney et al. (2009).

Method

Participants and Setting

A total of 4824 students (52% male, 48% female) in the second through sixth grades were included in the sample. The ethnicity breakdown of the sample was 94% White, not Hispanic; 2% Native American; 2% Asian or Pacific Islander; 1% Hispanic; and 1% Black, not Hispanic. Participants were enrolled in seven elementary schools, located within five school districts in the rural and suburban Midwest. The sample consisted of 31% students receiving free or reduced-price lunch, and 8% students receiving special education services.

As part of an ongoing benchmark assessment plan and response to intervention implementation (see Bollman, Silberglitt, & Gibbons, 2007, for additional information), all students in participating schools were assessed using CBM-R every fall, winter, and spring of every year across the grade levels studied. All students for whom relevant data were collected participated in the study. Data were gathered between the years of 2001 and 2005 for each student in the sample. Data from across multiple years were available for many of the students in the sample; thus many students participated in the study at multiple grade levels.

Measures

The CBM-R measures were developed or selected to be curriculum-independent, grade-level appropriate, and of equivalent difficulty across probes within each grade level. Each probe contained approximately 250

words typed on an 8 by 10 piece of paper in a grade-appropriate font (range = 12–14 point with larger print for the second and third grades). Probes were selected from a standardized set of commercially available published passages (AIMSweb; Howe & Shinn, 2002). The technical manual for AIMSweb passages (Howe & Shinn, 2002) and peer-reviewed research (Wayman et al., 2007) provide evidence that test–retest reliability and alternate-form reliabilities approximate or exceed .90. Criterion-related validity coefficients typically approximate the .70–.90 range. The same set of CBM-R probes were used across all districts for all 4 years of the study.

Procedure

CBM-R data were collected in 4-month intervals in the fall, winter, and spring of each academic year. Standardized administration and scoring procedures were implemented for each administration, which was conducted by trained school-based personnel and high school honors students. These individuals were trained to administer three successive CBM-R probes and score WRCM using standardized procedures that have appeared in the published literature (Howe & Shinn, 2002; Shinn, 1988, 1989). Administration occurred outside of the classroom to minimize distractions, in a separate room or quiet hallway while the administrator and student were seated at desks or in chairs.

Training included a brief instructional session that lasted approximately 1 hr. Each training session was followed by an assessment of competency that included practice and evaluation of scoring procedures. No formal criterion for scoring was established until the final years of the study. In 2003–2004 and 2004–2005, each scorer was required to come within 2 WCPM of the correct score on three consecutive videotaped assessments. Although data were not maintained, few scorers required more than one opportunity to meet criterion. Finally, inter-rater reliability was not evaluated; however, CBM-R procedures have established high levels of inter-rater reliability that appear consistently within the published literature. A survey of 10 arbitrarily selected CBM-R studies that were published between 1995 and 2005 reported inter-rater reliabilities that were consistently above .95.

Analytical Plan

A LMM analysis was performed to investigate evidence of seasonal effects for CBM-R. The present LMM analysis considered two hypothesized growth models: linear and piecewise (Duncan et al., 2006; Flora, 2008). The linear model assumes a static growth rate across time points—meaning that regardless of change over time, the estimated growth rate from the linear model is identical to the sample's growth rate. In contrast, the piecewise model assumes that the developmental trajectory is not consistent across time, making it possible to provide two different linear trends. Thus, the piecewise model takes into account a possible nonlinear trajectory for CBM-R. Based on the three CBM-R data points used in this study, the piecewise model can be parameterized into two levels of analyses. First, Level 1 can be expressed as

$$y_{it} = \beta_{0i} + \beta_{1i}\lambda_{1t} + \beta_{2i}\lambda_{2t} + \varepsilon_{it} \quad (1)$$

where y_{it} is the outcome of CBM-R for subject i at time t, β_{0i} is the value of CBM-R for subject i in fall, β_{1i} is the first linear trend of subject i from fall to winter, λ_{1t} is the value of the linear trend at time t (Week 0 and Week 18), β_{2i} is the second linear trend of subject i from winter to spring, λ_{2t} is the value of the linear trend at time t (Week 18 and Week 36), and ε_{it} is unique measurement error for subject i at time t.

Level 2 can be described as

$$\beta_{0i} = \beta_0\zeta_{0i} \quad (2)$$

$$\beta_{1i} = \beta_1\zeta_{1i} \quad (3)$$

$$\beta_{2i} = \beta_2\zeta_{2i} \quad (4)$$

where β_0 represents the average initial status for all subjects, β_1 refers to the average of the first linear growth rate for all subjects, and β_2

refers to the average of the second slope trend for all subjects. ζ_{0i}, ζ_{1i}, and ζ_{2i} refer to the random deviation of the initial status, the first growth rate, and the second growth rate, respectively. As a result, the piecewise model presented earlier assumes that the second slope is statistically different from the first slope, indicating that there is a seasonal effect for CBM-R throughout the academic year.

The first step in this analysis was to compare the two models (linear and piecewise models) using fit indices, which included the likelihood ratio test (LRT), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The three fit indices are considered to be "goodness-of-fit" indices, where the smallest fit statistics are viewed as the best model. If results from these goodness-of-fit indices reveal that the linear model is better than the piecewise model, then the slope is the same from fall to spring and there is no evidence of a significant seasonal effect for CBM-R across the academic year. Conversely, if goodness-of-fit indices are better for the piecewise model than the linear model, then we conclude that the slopes are different from fall to winter and from winter to spring and there is evidence of a seasonal effect across the academic year. If the piecewise model is the better fit, then an additional model is needed to estimate the extent to which the first slope differs from the second slope.

The additional model, which is called *added growth model,* would be identical to the piecewise model described earlier, except for the supplementary time factor loading matrix (Flora, 2008). For example, in the piecewise model, the time factor loading can be described as

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 18 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (5)$$

where the first column is set at 1, meaning that the outcome of the fall CBM-R score (intercept) does not change across time. The second column refers to the first linear time metric indicating change in CBM-R score from fall to winter (Flora, 2008). Similarly, the third column represents change in CBM-R score from winter to fall.

For the added growth model, the time factor loading can be expressed as:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 36 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (6)$$

The matrix in Equation 6 is identical to the matrix in Equation 5, with the exception of the second column with 0, 18, and 36. The different time factor loadings for the additional model is applied to examine the extent to which linear change in CBM-R from fall to winter increases or decreases compared to that from winter to spring. In other words, the added growth model analysis makes it possible to contrast the first linear slope and the second linear slope. The results from the additional analysis can also be viewed as an effect size estimate between the first and second linear slope (Flora, 2008).

Data Preparation

Overall, the rates of missing data for fall, winter, and spring seasons were less than 4% for all grade levels. All LMM analyses were conducted using the SAS MIXED program with full maximum likelihood estimation that automatically conducts all adjustments for missing data. Monthly CBM data used in this study were nested within districts. Therefore, a three-level LMM was used to satisfy an assumption that data collected in this study might not be independent among school districts. An alpha level for all statistical tests in this study was set at .001.

Model assumptions were evaluated, which included a test of normality using Kolmogorov-Smirnove statistics, skew, and kurtosis, and homoscedasticity of residuals by plotting raw residuals at each level (Levels 1 and 2). No violations were detected. The Level 3 LMM analysis was used to test whether CBM data used in the study were nested within school districts. As a result, variations in the slopes and intercepts between school districts were not significantly different

Table 1
Descriptive Statistics for the Level and Change of CBM-R

	Seasonal Benchmark Levels			Weekly Change	
Population	Fall *M* (*SD*)	Winter *M* (*SD*)	Spring *M* (*SD*)	Fall Season (Fall–Winter) *M* (*SD*)	Spring Season (Winter–Spring) *M* (*SD*)
General education					
Second	57 (32)	88 (34)	107 (35)	1.71 (.73)	1.02 (.64)
Third	82 (37)	107 (38)	125 (40)	1.38 (.79)	0.97 (.75)
Fourth	106 (38)	128 (39)	141 (41)	1.19 (.72)	0.78 (.70)
Fifth	125 (40)	145 (41)	160 (42)	1.12 (.76)	0.83 (.75)
Sixth	142 (41)	158 (42)	171 (43)	0.89 (.71)	0.74 (.75)
Total	99 (47)	122 (45)	137 (46)	1.26 (.74)	0.87 (.72)
Special education					
Second	33 (28)	55 (35)	74 (38)	1.17 (.81)	1.08 (.63)
Third	52 (33)	72 (39)	88 (41)	1.09 (.81)	0.89 (.78)
Fourth	77 (40)	96 (43)	108 (45)	0.99 (.72)	0.67 (.65)
Fifth	89 (43)	106 (45)	118 (47)	0.90 (.76)	0.67 (.79)
Sixth	106 (47)	119 (50)	132 (51)	0.70 (.72)	0.73 (.69)
Total	72 (45)	89 (48)	104 (49)	0.97 (.77)	0.81 (.71)

Note. CBM-R = curriculum-based measurement of oral reading. Skew and kurtosis were within reasonable limits ($z \leq 2.00$) for all conditions. *N* values for second, third, fourth, fifth, and sixth grades were, respectively, 1790, 1885, 1913, 1942, and 1144 for general education; 111, 134, 170, 192, and 116 for special education.

from zero for all grade levels, meaning that both the slopes and intercepts at the district levels were identical. In addition, such a result supports evidence that the Level 2 LMM analysis introduced in the method section is appropriate for this study.

Results

Summaries of the means and standard deviations for CBM-R at each time point are presented in Table 1. As shown in Table 1, the average growth as measured with CBM-R was positive for all conditions across seasons (fall–winter, winter–spring), grade levels (second to sixth) and populations (general education, special education).

Comparison Between Linear Model and Piecewise Model

Results for the comparisons of linear and piecewise models are presented in Table 2. For students in general education, the piecewise model provided the best fit for tri-annual assessment data, as evidenced by the magnitudes of LRT, AIC, and BIC observed for the linear and piecewise models respectively. The χ^2 difference test between the two models also revealed that the piecewise model fit significantly better than the linear model.

For students in special education, on the basis of the LRT criterion, the piecewise model for all grade levels was viewed as the best-fitting model. Significant differences in LRTs between the two models, however, were found in the second through fourth grades. There were inconsistencies with respect to AIC and BIC criteria, which might indicate the linear model fit better in some cases within the sample of special education students. Although the results of AIC and BIC are noted, the LRT was the primary test/criterion used here.

Table 2
Deviation Between Linear and Piecewise Models for General Education Students and Special Education Students

	Linear Model			Piecewise Model			Differences (Linear-Piecewise Model)		
Grade	LRT	AIC	BIC	LRT	AIC	BIC	LRT	AIC	BIC
General education									
Second	46231.7	46239.7	46261.7	45542.6	45556.6	45595.0	689.1*	683.1	666.7
Third	49566.9	49574.9	49597.0	49357.7	49371.7	49410.5	209.2*	203.2	186.5
Fourth	49768.1	49776.1	49798.3	49520.1	49534.1	49573.0	248.0*	242.0	225.3
Fifth	50974.9	50982.9	51005.2	50866.6	50880.6	50919.6	108.3*	102.3	85.6
Sixth	29892.9	29900.9	29921.1	29873.6	29887.6	29922.9	19.3*	13.3	−1.8
Special education									
Second	2816.3	2824.3	2835.1	2802.6	2816.6	2835.6	13.7*	7.7	−0.5
Third	3501.8	3507.8	3516.5	3487.9	3495.9	3507.5	13.9*	11.9	9.0
Fourth	4394.9	4402.9	4415.4	4378.9	4392.9	4414.9	16.0*	10.0	0.5
Fifth	5090.6	5098.6	5111.6	5083.3	5097.3	5120.1	7.3	1.3	−8.5
Sixth	3048.2	3056.2	3067.2	3046.3	3060.3	3079.5	1.9	−4.1	−12.3

Note. LRT = likelihood ratio test; AIC = Akaike information criterion; BIC = Bayesian information criterion.
*$p < .001$.

Comparison Between the First Slope and Second Slope for General and Special Education

Table 3 presents estimates of both the first (fall–winter) and second (winter–spring) linear slope. In the cases of both students in general education and special education, all of the seasonal linear slopes were positive with magnitudes significantly different than zero ($p < .001$). That observation was consistent with the conclusion that reading achievement accelerated continuously throughout the academic year for all student samples. Although the mean rate of growth across seasons was consistently greater for students in general education as compared with students in special education, it was the first slopes that were most distinct. Further, the differences in the second slopes varied by grade level.

Additional analysis was conducted to examine whether the first linear slopes (fall) were significantly different from the second slopes (spring). Table 3 presents the differences between the two linear slopes along with the results of significance tests. Figure 1 depicts the slopes and model differences.

In the sample of general education students, statistically significant differences ($p < .001$) were observed between the first and second slopes in all grade-level conditions (Table 3). That is, the rate of growth for all grade levels dropped in the second part of the academic year. It is relevant to note that the magnitude of difference between the two slopes (seasonal effects) and standardized effect sizes consistently decreased as grade level increased. This might indicate that the seasonal effect in reading is larger in the lower primary grades than the upper primary grades. The power estimates for detecting seasonal effects for students in general education were very large for all grade levels (range = 0.96–1.00). This indicates a good chance of detecting a false null hypothesis (linear model) when the alternative model (piecewise model) is correct (Duncan et al., 2006).

In the sample of special education students, the average of the first slope (fall) was

Table 3
Comparison Between the First Slope and the Second Slope for General Education Students and Special Education Students

Grade	Intercept $\hat{\beta}_0$ (*SE*)	Fall/First Slope (Fall to Winter) $\hat{\beta}_1$ (*SE*)	Spring/ Second Slope (Winter to Spring) $\hat{\beta}_2$ (*SE*)	Difference Between Fall Slope and Spring Slopes $\hat{\beta}_2{}^a$ (*SE*)	Standardized Effect Sizes[b]	Power
General education						
Second	57.42* (.76)	1.71* (.02)	1.02* (.01)	0.69* (.02)	34.50	1.00
Third	82.31* (.85)	1.38* (.02)	0.97* (.02)	0.41* (.03)	13.60	1.00
Fourth	105.78* (.93)	1.19* (.02)	0.78* (.02)	0.41* (.12)	3.41	1.00
Fifth	124.43* (.89)	1.12* (.02)	0.83* (.02)	0.29* (.12)	2.41	1.00
Sixth	142.49* (1.21)	0.88* (.02)	0.74* (.02)	0.14* (.16)	0.87	0.96
Special education						
Second	34.02* (2.67)	1.17* (.08)	1.08* (.06)	0.09 (.10)	0.90	0.86
Third	50.85* (2.81)	1.09* (.07)	0.89* (.06)	0.20 (.10)	2.00	0.87
Fourth	75.50* (2.99)	0.99* (.05)	0.67* (.05)	0.32* (.08)	4.00	0.92
Fifth	86.28* (3.06)	0.90* (.05)	0.67* (.06)	0.23 (.09)	2.50	0.54
Sixth	106.98* (4.40)	0.69* (.07)	0.73* (.06)	−0.04 (.10)	0.40	0.11

Note. SE = standard error.
[a]The estimated coefficients of the second slope from an added growth model.
[b]Standardized effect sizes = $\hat{\beta}_2$/SE, which is a strategy developed by Satorra and Saris (1985), was used to calculate power.
*$p < .001$.

greater than the second slope (spring) in all grades except for sixth; however, those differences were not statistically significant except in the fourth-grade condition (Table 3). Moreover, the pattern of reduced seasonal effects for progressively higher grades did not generalize to the sample of special education students. Instead, the magnitude was greatest in fourth grade (difference in slopes = 0.32, $p < .001$), somewhat smaller in third (0.20, $p > .001$) and fifth (0.23, $p > .001$), and hardly perceptible in second (0.09, $p > .001$) and sixth grades (−0.04, $p < .001$). The range of standardized effect sizes and power was variable across grades. Based on the results, it appears that, in general, the observed slopes in the sample of special education students were relatively consistent across seasons—in all conditions but fourth grade.

Discussion

Many of the published estimates of CBM-R annual growth rates rely on the assumption that the pattern is consistent and linear throughout the academic year (Deno et al., 2001; Fuchs et al., 1993). Although there is evidence for seasonal effects within some studies (Ardoin & Christ, 2009; Fuchs et al., 1993; Graney et al., 2009), the phenomenon of seasonal effects are generally ignored within the professional literature. This study yielded (a) estimates for the magnitude CBM-R annual growth along with those within the fall and spring seasons, and estimates were derived for both general education and special education populations; (b) both linear and piecewise growth models; and (c) hypothesis tests to examine whether fall growth is typically more substantial than spring growth.

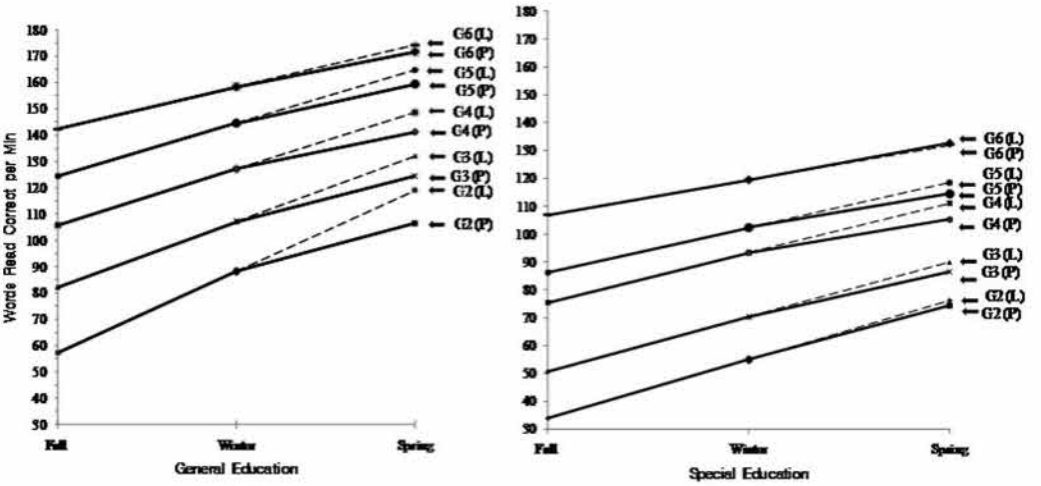


Figure 1. Differences for linear (L, solid line) and piecewise (P, dashed line) models of growth (G) for students in second through sixth grade within either the general (left) or special education population (right)

CBM-R growth was modeled and analyzed for five primary grades (second through sixth) and two student populations within and across seasons.

The results of this study support four general conclusions. Descriptive data make clear that, first, annual growth is greater among students in the general education population than for those in the special education population, especially in the fall season; and, second, more growth occurs in the early grades with less in the upper grades. Modeling and inferential analysis provide support for the third and fourth conclusions that follow. That is (third conclusion), there is a significant *seasonal effect* such that more growth occurs in the fall season than in the spring season for general education, but the magnitude of that effect declines with each progressive grade level. Finally, the seasonal effect is less pronounced among students in the special education population. In general, linear growth should not be assumed, seasonal effects should be expected—at least within the general education population—and piecewise models tend to fit better than linear models of growth.

Grade and Population Effects

Larger magnitudes of growth were observed for students in general education as compared to special education. This outcome was observed across five districts where CBM-R was used to screen, monitor, and evaluate instructional effects as part of a problem solving model. The systematic and ongoing use of CBM-R for these purposes did not establish equivalent growth rates between the student populations. Despite inconsistent growth rates across student groups, the rate of growth among students in the special education population was greater than observed in other studies (Deno et al., 2001; Fuchs et al., 1993). For example, the most recent estimates for average growth among students in the special education population were ~0.60 WRCM/week (Deno et al., 2001). That rate of growth was relatively consistent across grades. In contrast, the average growth estimates among students in the sample of students in special education used in this study were within the range of 0.67–1.17 WRCM/week (Table 3). Although cross-study comparisons cannot be used to establish that differences in systems-level practices (e.g., problem solving) and instructional practices were causally related to

the observed differences in growth rates, future research should investigate the phenomena more closely.

The results of this study converge with previous research (Deno et al., 2001; Fuchs et al., 1993) to support the conclusion that there is more CBM-R growth within the early primary grades. This phenomenon was consistent for both fall and spring seasons in both the general education and special education populations (Table 1). There was a steady decline for both general education and special education in weekly growth as grade level was increased. There was a corresponding decline in the magnitude of seasonal effects as grade levels increased. Results also support the conclusion that seasonal effects are more robust for students served within general education and less robust for students served within special education.

Seasonal Effect

The primary purpose of this research was to determine whether the rate of CBM-R growth is consistent across the year. Notwithstanding limitations, a seasonal effect was sufficiently robust and consistent across grades and student populations to warrant serious consideration. Estimates of weekly CBM-R growth within season are reported in Table 1. They are also depicted in Figure 1. Both visual analysis and statistical analysis of growth models support the conclusion that there was a *difference in the average rate of growth across seasons*, and that difference occurred in all grades, but was generally more robust in the lower grades of general education (Figure 1). With few exceptions, more growth occurred in the fall than in the spring, which is depicted as the differences in the slopes within Table 3.

The *magnitudes of seasonal effects were inconsistent* across grades and student populations. In general education, the difference between fall and spring slopes ranged from 0.14 to 0.69 per week, and all the differences were statistically significant with substantial effect sizes (Table 3). That effect is illustrated in Figure 1, which reveals that the interactions between seasonal effect and student population were most distinct in second grade. That is, the seasonal effect was most robust for the sample of students in second-grade general education classes, with a difference in growth from fall to winter of 0.69, and minor in comparison for students in second-grade special education classes, with a difference of 0.09.

The results of this study support the conclusion that growth differs between fall and spring seasons, which is inconsistent with a number of currently held views in the field. The consistency of the results across grades for general education begs the question of why this effect may be occur and what are the potential implications. There are a number of alternate hypotheses that offer some insight on the phenomenon and directions for future research.

Explanatory Hypotheses

This study was designed to examine *whether* seasonal effects are likely to occur in cases of CBM-R tri-annual assessment and challenge the assumption of linear growth. Indeed, the results do provide support for seasonal effects; however, this study was not designed to examine the cause of seasonal effects. That is left to future research to isolate the cause(s) and potential remedies to accelerate reading achievement within seasons with deficit growth. The result of both this study and Ardoin and Christ (2008) converge to provide evidence that the seasonal effect is likely to coincide with deficit levels of growth in the latter part of the academic year, or the spring season, especially for the students in the early primary general education population. It seems that changes in instructional conditions, or student motivation, might explain deficit rates of growth within the spring season unless there is an active intervention to prevent those changes. It seems that data-based decision making could function to reduce, or reverse, the pattern of seasonal effects that were replicated across this study and that of Ardoin and Christ.

Teacher expectations. There are a number of ways in which teacher expectations can have an influence on student behavior,

which might be conveyed through discrepant use of praise and punishment statements, content-relevant feedback, individualized attention, differentiated instruction, and variability in the opportunities to respond and engage in instructionally relevant practice (Good, 1981). Prior research yielded conclusions that there are indeed sometimes differences in teacher treatment of high- and low-achieving students over the course of the year, which affected the amount and type of instruction and feedback that students received at different points in the academic year (Good et al., 1980). Researchers concluded that teachers are more invested in influencing and shaping student behavior at the beginning of the school year than later in the year. Furthermore, the differentiation in teacher behavior may have sustained effects and become more influential as the school year progresses, resulting in lower achievement and growth later in the school year in low-achieving students. Other research also demonstrated the potential influence of teacher expectations, including expectation biases across ethnic groups (McKown & Weinstein, 2007) and ability groups (Rosenthal & Jacobson, 1966).

Classroom management. There is evidence that classroom management and organizational strategies are implemented with greater fidelity and effort in the early part of the academic year (Cameron, Connor, & Morrison, 2004). The same research provides evidence that that those educators who approach classroom management and organization with the greatest fervor early in the year are those same educators whose practices drop off most substantially later in the academic year. Improved levels of engagement and instructional gain are clearly associated with good classroom management and organization, and are among better predictors of student achievement (Brophy, 1987). Taken together, research provides a foundation to postulate that differential levels of classroom management and organization across seasons might influence student achievement. This effect is hypothesized to be the greatest at lower grade levels, as younger students require greater direction in organization practices (Cameron et al., 2004). The pattern of seasonal effects observed in this study coincides with a classroom management and organizational hypothesis.

Intervention to accelerate seasonal growth. The results of this study indicate that the seasonal effect with deficit growth in the spring season was most pronounced in the middle primary grades (second and third) and were minimized for the special education population. The results of Graney et al. (2009) indicate that the seasonal effect can function so that there is deficit growth in the fall season and accelerated growth in the spring. As noted in the Introduction, the data set that was analyzed by Graney et al. (2009) derived from a university and school partnership to implement a data-based decision making and evidence-based interventions. Those factors might have influenced the patterns of seasonal effects. There is a substantial literature base to support the conclusion that data-based decision making can function to accelerate academic growth (Carnine & Granzin, 2001; Fuchs & Fuchs, 1986; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Fuchs, Fuchs, Hamlett, & Whinnery, 1991; Graney et al., 2009), and there is substantial theoretical support for the use of evidence-based practices, which were both components of service delivery that were targeted by the partnership program. It might be that progress monitoring and data use accelerated learning and ameliorated the deficit growth associated with seasonal effects. Indeed, students who are served within special education have explicit goals and their progress is monitored. Although it is not possible to draw any final conclusions, it may well be that the lack of convergence in seasonal patterns between Graney et al. and other studies—along with minimal seasonal effects observed for the special education population—provides further evidence for the use of data-based decision making to improve growth during otherwise deficit periods.

Limitations and Future Research

The sample studied was entirely from five relatively ethnically homogeneous dis-

tricts in a single midwestern state. Future research should attempt to corroborate these findings across a more diverse sample of students and educational programs. Finally, the results of this study derive from data collected at three points in the academic year (fall, winter, spring). Future research should examine whether the same conclusions would result from the analysis of a more robust data set (e.g., weekly, monthly) and with more precise dates that define the data collection schedule. Graney et al. (2009) correctly criticize the methods of this study and that of Ardoin and Christ (2008) because the precise dates of data collection were not used in the analysis; rather, both studies rely on data that are centered by month and not day. Although the authors of this and the prior study believe the conclusions would be substantially similar with different centering methods or more precise dates, the criticism ought to be accounted for in future research. Future research should investigate the nature of growth across summative and formative measurement conditions. Research should also examine the effect on educational decisions when the potential for seasonal effects is ignored and consistent annual growth is assumed.

Implications for School-Based Decision Making

CBM-R was developed to guide routine instructional decisions and enhance instructional effects (Deno, 1986), and it has emerged as one of the primary methods of assessment to guide problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention models of service delivery (National Association of State Directors of Education, 2005). As described in the Introduction of this study, CBM-R is used to establish expectations for both the level and rate of reading achievement throughout the primary grades; therefore, the results of this study provide at least two clear implications. First, there is good reason to question the validity of linear growth estimates as criteria to guide progress monitoring (i.e., discrepant in rate) and dual-discrepancy decisions (i.e., discrepant in both level and rate; Fuchs, Fuchs, McMaster & Al Otaiba, 2003). Second, seasonal effects can be either admired as a phenomenon or targeted for intervention. These implications are addressed in what follows.

First, if CBM-R estimates of growth are used to evaluate instructional effects and response to intervention, then those who interpret data should be aware that the magnitude of growth might be less in the spring than in the fall. There are two weighty implications that derive from this study. First, a dual-discrepancy approach to special education eligibility, which compares the observed level and rate of achievement to criterion standards (Fuchs, Fuchs, McMaster & Al Otaiba, 2003), might function to overidentify students during deficit seasons of growth and underidentify students during accelerated seasons of growth. If the data follow the pattern observed in this study, then it is likely that the dual-discrepancy model might underidentify students in the fall season and overidentify students in the spring season. That consequence is counter to the values of early intervention and prevention inherent within problem solving and response to intervention. Researchers and practitioners should pay close attention to the patterns of growth and the potential for seasonal effects to influence decisions at the local level. Accelerated rates of growth should be expected in the fall.

Second, the pattern of seasonal effects that are observed across populations and studies provide preliminary evidence that goal setting, progress monitoring and data-based decision making can minimize, eliminate or reverse the pattern of seasonal effects (e.g., Graney et al., 2009). Although future research is necessary, it is a working thesis that the systematic use of data to evaluate instructional effects and select effective instructional strategies effectively accelerates learning and prevents problems at the individual, group, and systems level. After all, data-based decision making is a central tenet of problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention (National Association of State Directors of Education, 2005); however, data must be both systematically collected *and* sys-

tematically used to establish effects (Fuchs et al., 1992, 1991; Fuchs & Fuchs, 1986). The mere act of data collection without data use typically fails to establish effects (Fuchs & Fuchs, 1986). The explanatory hypotheses proposed in the previous section establish that deficit growth in the spring season might be remediated at the systems level if data are systematically collected and systematically used. If detected, this problem (seasonal effect) is ripe for problem analysis at the local level. Deficit growth within a single season should not be admired as an acceptable phenomenon, but, instead, it should be targeted for analysis and intervention.

Christ (2008) defined problem analysis as the "collection, summary, and use of information to systematically test, reject, or verify relevant hypotheses to establish problem solutions" (p. 159). That is, deficit rates of growth in the spring season might be remediated if expectations remain high and data are used to evaluate and maintain high standards for instructional effects. There is no reason to assume that the seasonal effect is inherent to the educational system. Instead, hypotheses related to the causal and maintaining variables should be developed—such as those presented earlier—and tested through intervention to remediate the problem. It is likely that the pattern of seasonal effects is causally related to what occurs in the classroom; and what occurs in the classroom can be influenced by the systematic used of data to evaluate of response to intervention for universal, supplemental, and intensive tiers of service delivery. This study provides evidence of seasonal effects as a potential target for intervention at the system, grade, or classroom level.

References

AIMSweb. (2002). *Reading curriculum-based measurement (R-CBM) standard oral reading fluency assessment passages.* Retrieved October 2003, from http://www.aimsweb.com

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum based measurement slope estimate using data from tri-annual universal screenings *School Psychology Review, 37,* 109–125.

Bollman, K. A., Silberglitt, B., & Gibbons, K. A. (2007). The SCRED Model: Incorporating systems-level organization and a multi-tiered problem-solving process for intervention delivery. In M. Burns, S. Jimerson, & A. VanDerHayden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 321–332). New York: Springer.

Brophy, J. (1987). Synthesis of research on strategies for motivating students to learn. *Educational Leadership, 45*(2), 40–48.

Cameron, C. E., Connor, C. M., & Morrison, F. J. (2004). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology, 43,* 61–85.

Carnine, D., & Granzin, A. (2001). Setting learning expectations for students with disabilities. *School Psychology Review, 30,* 466–472.

Christ, T. J. (2008). Best practices in problem analysis. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 159–176). Bethesda, MD: National Association of School Psychologists.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74,* 68.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219–232.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15,* 358–374.

Deno, S. L. (2002). Problem solving as best practices. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 37–56). Bethesda, MD: National Association of School Psychologists.

Deno, S. L. (2005). Problem solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 10–42). New York: Guilford Press.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30,* 507–524.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). An introduction to latent variable growth curve modeling: Concepts, issues, and applications (2nd ed.). Mahwah, NY: Lawrence Erlbaum Associates.

Flora, D. B. (2008). Specifying piecewise latent trajectory models for longitudinal data. *Structural Equation Modeling, 15,* 513–533.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53,* 3–14.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10,* 43–52.

Fuchs, L. S., Fuchs, D., & Stecker, P. M. (1989). Effects of curriculum-based measurement on teachers' instructional planning. *Journal of Learning Disabilities, 22,* 51–59.

Fuchs, L. S., & Shinn, M. R. (1989). Writing CBM IEP objectives. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 131–152). New York: Guildford Press.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58,* 436–450.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22,* 27–48.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Whinnery, K. (1991). Effects of goal line feedback on level, slope, and stability of performance within curriculum-based measurement. *Learning Disabilities Research and Practice, 6,* 66–74.

Fuchs, D., Fuchs, L. S., McMaster, K. N., & Al Otaiba, S. (2003). Identifying children at risk for reading failure: Curriculum-based measurement and the dual-discrepancy approach. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 431–449). New York: Guilford Press.

Good, T. L. (1981). Teacher expectations and student perceptions: A decade of research. *Educational Leadership, 38*(5), 415–422.

Good, T., Cooper, H., & Blakey, X. (1980). Classroom interaction as a function of teacher expectations, student sex and time of year. *Journal of Educational Psychology, 72*, 378–385.

Graney, S. B., Missall, K. N., Martinez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School* Psychology, 47, 121–142.

Howe, K. B., & Shinn, M. M. (2002). Standard *reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Retrieved October 2004, from http://www.aimsweb.com

Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education for Students Placed at Risk, 9,* 169–188.

McKown, C., Weinstein, R. S. (2007). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*, 235–261.

National Association of State Directors of Education. (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: Author.

National Reading Panel. (2000). *Report of the National Reading Panel Subgroups: Teaching children to read* (No. NIH 00–4754). Bethesda, MD: National Institute of Child Health and Human Development.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports, 19,* 115–118.

Satorra, A., & Saris, W. E. (1985). The power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50*, 83–90.

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education, 34,* 164–172.

Shinn, M. R. (1988). Development of curriculum-based local norms for use in special education decision-making. *School Psychology Review, 17,* 61–80.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children.* New York: Guildford Press.

Shinn, M. (2008). Best practices for using curriculum based measurement in a problem solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 243–262). Bethesda, MD: National Association of School Psychologists.

Stage, S. A. (2001). Program evaluation using hierarchical linear modeling with curriculum-based measurement reading probes. *School Psychology Quarterly, 16,* 91–112.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41,* 85–120.

Date Received: May 30, 2007
Date Accepted: April 21, 2010
Action Editor: Tanya Eckert ■

Theodore J. Christ, PhD, is an associate professor in the Educational Psychology Department at the University of Minnesota in Minneapolis—Saint Paul. He is the codirector of the Research Institute for Problem Solving. He is engaged in research and training to evaluate, develop, and enhance the tools and procedures that enable data-based decision making and problem solving within school settings. Interests include progress monitoring, curriculum-based measurement, direct behavior ratings, and computer adaptive testing.

Benjamin Silberglitt completed his PhD in Educational Psychology at the University of Minnesota at Minneapolis—Saint Paul. His areas of interest include student assessment school organization and system-level change, problem solving, response to intervention, and data-based decision making. He is director of Software Applications for TIES, a technology cooperative of 39 Minnesota school districts.

Seungsoo Yeo, PhD, is a researcher at Korean Educational Development Institute. He received his PhD from the University of Minnesota at Twin Cities in 2008. His primary research interests include curriculum-based measurement, response to intervention, and learning disabilities. His methodological work has focused on latent growth modeling and multilevel meta-analysis.

Damien C. Cormier, MA, is a school psychology doctoral student in the Educational Psychology Department at the University of Minnesota. He is a graduate research assistant at the National Center on Educational Outcomes. His research interests are in the areas of large-scale assessment policy and practice, measurement issues in individual assessment, high-functioning autism, and cultural and linguistic biases in assessment.

Copyright of School Psychology Review is the property of National Association of School Psychologists and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Karla Ramy

From:Karla Ramy
Sent:26 Aug 2013 15:35:58 +0000
To:What Works
Subject:Read Naturally - Tom Ihnot
Attachments:Christ Silberglitt Yeo Cormier (2010).pdf, Heistad.docx
Jill Constantine,

I am requesting that What Works Clearinghouse (WWC) change its Beginning Reading effectiveness rating on reading fluency for Read Naturally, Inc. from mixed effects to positive effects for several reasons. Please consider and respond to the following information.

First, in a recent WWC review of a study by Dr. David Heistad, statistically significant gains in oral reading fluency for third grade students across multiple schools were not considered in the evaluation of the effects of a Read Naturally intervention on fluency. After reviewing the most recent WWC Procedures and Standards Handbook (February, 2013), the reason for exclusion may be due to a simple misunderstanding.

Page 15 of the WWC Procedures and Standards Handbook (Outcome Eligibility and Reporting section) states the following:

"….outcome measures is that they not be over aligned with the intervention. When outcome measures are closely aligned with or tailored to the intervention, the study findings may not be an accurate indication of the effect of the intervention. For example, an outcome measure based on an assessment that relied on materials used in the intervention condition but not in the comparison condition…."

The measure of oral reading fluency in Dr. Heistad's study was *The Reading Fluency Monitor*. It should be noted that *The Reading Fluency Monitor* is not a component of the Read Naturally Intervention. The *Reading Fluency Monitor* was developed by RMC Research Corporation, located in Portsmouth, New Hampshire. Dr. Heistad's implication that *The Reading Fluency Monitor* was developed by Read Naturally, Inc. was a miscommunication in the Test Instruments section of his study. Heistad's verification of this misunderstanding is attached.

Second, the Arvans study summary analysis underestimated the effect of the Read Naturally group on oral reading fluency. The Read Naturally group had a large effect size of .81 for fluency. The control EAU group had a moderate effect size of .57 for fluency. Both groups had excellent fluency gains. It should be noted that the use of the Read Naturally Intervention was extensive at the school before the study was conducted. The teachers were well versed in the Read Naturally strategy. Students used Read Naturally in the first and second grade in the two years prior to the study. Also, some students in the EAU group used Read Naturally during the study. That is, it seems as though the EAU group may not have been demonstrative of a true control group. The difference between the Read Naturally group's .81 effect size and EAU's .57 effect size was .24 after eight weeks. An effect size difference of .24 is significant when extrapolated over a school year.

Ethan R. Van Norman, M.A., performed a similar analysis to what appeared in the Christ and Davie study (2008). In that study, the authors first calculated a slope estimate from three time points for each student in the EAU and the experimental group. The slope estimate represented the number of words read correct per minute (WCPM) improvement per week. The mean and standard deviation of slopes

were then calculated for each group. The percent of improvement of the experimental group in relation to the EAU group was calculated. After this, the authors used the percent of improvement and applied it to an aggressive rate of growth (1.50 WCPM improvement per week). That value was then multiplied by 36, and 1.50 was multiplied by 36. The difference between those values was then interpreted as a hypothetical effect if the Read Naturally Intervention was delivered across an entire school year.

A similar analysis was conducted on the Arvans dataset. Slope estimates were calculated for each student from two observations eight weeks apart. The mean slope value for the experimental group was 2.92 WCPM improvement per week (SD = 1.54) compared to the EAU group, which had a mean slope estimate of 2.24 (SD = 2.36). The .68 difference in mean slope for the Read Naturally Intervention group represents a 30% improvement over the EAU group. Assuming an aggressive rate of growth of 1.50 WCPM for typical students, a 30% increase would translate to a 1.95 rate of growth. Extended across 36 weeks, this represents a net increase of 70 WCPM compared to 54 WCPM with a typical rate of growth (1.50 WCPM improvement per week across 36 weeks).

Although not ideal, slope estimates from two time points have been used to summarize growth in previous CBM-R research studies (e.g., Christ, Silberglit, Yeo & Cormier, 2010). See the attached document for further details.

Another way to analyze the Arvans study is to use the Hasbrouck-Tindal Oral Reading Fluency Norms (http://www.readnaturally.com/howto/orftable.htm). The table suggests that third grade students at the 50th percentile have an average weekly improvement of 1.1 WCPM. The performance of the Read Naturally group in the Arvans study is also impressive when compared to the typical rate of fluency growth of third grade students.

Ethan R. Van Norman, M.A., did an analysis to extrapolate growth in the same manner using normative values. For third grade students, weekly growth estimates for students in the 50th percentile typically approximate 1.10 WCPM per week. Assuming the Read Naturally group's 30% improvement over the EAU group and the intervention is delivered for 36 weeks, a student in the 50th percentile, on average would be improving at a rate of 1.43 WCPM per week. After 36 weeks this would translate to an improvement of 51 WCPM compared to 40 WCPM if the student did not receive the intervention. This is a substantial difference.

Considering all of this information, Read Naturally programs should receive a positive effects rating for fluency. A timely reevaluation is critical so that WWC may change Read Naturally's effectiveness rating for fluency from mixed effects to positive effects. Please respond to the points I've raised in this email and let me know if you need further information.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally®
CELEBRATING
20 YEARS

I was mistaken when I indicated that the Reading Fluency Monitor was developed by Read Naturally. It was actually developed by the RMC Research Corporation out of Portsmouth, New Hampshire.

Sorry for the misunderstanding,

- Dave

David Heistad, Ph.D., Executive Director
Research and Evaluation Department
Bloomington Public Schools - ISD 271
1350 W. 106th Street
Bloomington, MN 55431
ph: 952-681-6485
cell: 952-250-5971

WhatWorks

From:WhatWorks
Sent:28 Aug 2013 13:40:22 +0000
To:'kramy@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot
Dear Mr. Ihnot,

Thank you for your email. The QRT will send a separate response to your email of August 26, 2013.

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Monday, August 26, 2013 11:36 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Jill Constantine,

I am requesting that What Works Clearinghouse (WWC) change its Beginning Reading effectiveness rating on reading fluency for Read Naturally, Inc. from mixed effects to positive effects for several reasons. Please consider and respond to the following information.

First, in a recent WWC review of a study by Dr. David Heistad, statistically significant gains in oral reading fluency for third grade students across multiple schools were not considered in the evaluation of the effects of a Read Naturally intervention on fluency. After reviewing the most recent WWC Procedures and Standards Handbook (February, 2013), the reason for exclusion may be due to a simple misunderstanding.

Page 15 of the WWC Procedures and Standards Handbook (Outcome Eligibility and Reporting section) states the following:

"….outcome measures is that they not be over aligned with the intervention. When outcome measures are closely aligned with or tailored to the intervention, the study findings may not be an accurate indication of the effect of the intervention. For example, an outcome measure based on an assessment that relied on materials used in the intervention condition but not in the comparison condition…."

The measure of oral reading fluency in Dr. Heistad's study was *The Reading Fluency Monitor*. It should be noted that *The Reading Fluency Monitor* is not a component of the Read Naturally Intervention. The *Reading Fluency Monitor* was developed by RMC Research Corporation, located in Portsmouth, New Hampshire. Dr. Heistad's implication that *The Reading Fluency Monitor* was developed by Read Naturally, Inc. was a miscommunication in the Test Instruments section of his study. Heistad's verification of this misunderstanding is attached.

Second, the Arvans study summary analysis underestimated the effect of the Read Naturally group on oral reading fluency. The Read Naturally group had a large effect size of .81 for fluency. The control EAU group had a moderate effect size of .57 for fluency. Both groups had excellent fluency gains. It should be noted that the use of the Read Naturally Intervention was extensive at the school before the study was conducted. The teachers were well versed in the Read Naturally strategy. Students used Read

Naturally in the first and second grade in the two years prior to the study. Also, some students in the EAU group used Read Naturally during the study. That is, it seems as though the EAU group may not have been demonstrative of a true control group. The difference between the Read Naturally group's .81 effect size and EAU's .57 effect size was .24 after eight weeks. An effect size difference of .24 is significant when extrapolated over a school year.

Ethan R. Van Norman, M.A., performed a similar analysis to what appeared in the Christ and Davie study (2008). In that study, the authors first calculated a slope estimate from three time points for each student in the EAU and the experimental group. The slope estimate represented the number of words read correct per minute (WCPM) improvement per week. The mean and standard deviation of slopes were then calculated for each group. The percent of improvement of the experimental group in relation to the EAU group was calculated. After this, the authors used the percent of improvement and applied it to an aggressive rate of growth (1.50 WCPM improvement per week). That value was then multiplied by 36, and 1.50 was multiplied by 36. The difference between those values was then interpreted as a hypothetical effect if the Read Naturally Intervention was delivered across an entire school year.

A similar analysis was conducted on the Arvans dataset. Slope estimates were calculated for each student from two observations eight weeks apart. The mean slope value for the experimental group was 2.92 WCPM improvement per week (SD = 1.54) compared to the EAU group, which had a mean slope estimate of 2.24 (SD = 2.36). The .68 difference in mean slope for the Read Naturally Intervention group represents a 30% improvement over the EAU group. Assuming an aggressive rate of growth of 1.50 WCPM for typical students, a 30% increase would translate to a 1.95 rate of growth. Extended across 36 weeks, this represents a net increase of 70 WCPM compared to 54 WCPM with a typical rate of growth (1.50 WCPM improvement per week across 36 weeks).

Although not ideal, slope estimates from two time points have been used to summarize growth in previous CBM-R research studies (e.g., Christ, Silberglit, Yeo & Cormier, 2010). See the attached document for further details.

Another way to analyze the Arvans study is to use the Hasbrouck-Tindal Oral Reading Fluency Norms (http://www.readnaturally.com/howto/orftable.htm). The table suggests that third grade students at the 50th percentile have an average weekly improvement of 1.1 WCPM. The performance of the Read Naturally group in the Arvans study is also impressive when compared to the typical rate of fluency growth of third grade students.

Ethan R. Van Norman, M.A., did an analysis to extrapolate growth in the same manner using normative values. For third grade students, weekly growth estimates for students in the 50th percentile typically approximate 1.10 WCPM per week. Assuming the Read Naturally group's 30% improvement over the EAU group and the intervention is delivered for 36 weeks, a student in the 50th percentile, on average would be improving at a rate of 1.43 WCPM per week. After 36 weeks this would translate to an improvement of 51 WCPM compared to 40 WCPM if the student did not receive the intervention. This is a substantial difference.

Considering all of this information, Read Naturally programs should receive a positive effects rating for fluency. A timely reevaluation is critical so that WWC may change Read Naturally's effectiveness rating for fluency from mixed effects to positive effects. Please respond to the points I've raised in this email and let me know if you need further information.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Karla Ramy

From:Karla Ramy
Sent:29 Aug 2013 14:50:33 +0000
To:What Works
Subject:Tom Ihnot
Jill Constantine,

I appreciate that What Works Clearinghouse will change the Beginning Reading findings on Read Naturally by revising the order of findings listed in the text on the front page of the report, the associated table and the website summary for the three reports with finding in more than one domain. Since educators are heading back to school, I hope the change can be made as soon as possible.

Also, I am anxious to have my request (email dated 8/26/13) to change WWC Beginning Reading finding on Read Naturally from *mixed effects on fluency* to *positive effects evaluated* as soon as possible.

Thank you,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Intervention Summary for *Read Naturally®*

Please review the intervention description below, and tell us whether we have described the intervention accurately, and whether we have included all relevant details. The information you provide may be incorporated into reports on *Read Naturally®* and posted to the WWC website for public access.

Program Description

The Read Naturally® program is a supplemental reading program that aims to improve the reading fluency, accuracy, and comprehension of elementary and middle school students using a combination of texts, audio CDs, and computer software. The program uses one of four products that share a common fluency-building strategy: Read Naturally® Masters Edition, Read Naturally® Encore, Read Naturally® Software Edition, and Read Naturally® Live. The common strategy includes: modeling of story reading, repeated reading of text for developing oral reading fluency, and systematic monitoring of student progress by teachers and the students themselves. Students work at their own reading level, progress through the program at their own rate, and work (for the most part) on an independent basis. The program can be delivered in three ways: (1) students use audio CDs with hard-copy reading materials (Read Naturally® Masters Edition, Read Naturally® Encore), (2) students use the computer-based version (Read Naturally® Software Edition), or (3) students use the web-based version (Read Naturally® Live). This intervention report includes studies of Read Naturally® Masters Edition and Read Naturally® Software Edition.

Background

Developed by Candyce Ihnot, the four Read Naturally® products are distributed by Read Naturally, Inc. Address: 2945 Lone Oak Drive, Suite #190, Saint Paul, MN 55121. Email: info@readnaturally.com. Web: www.readnaturally.com. Telephone: (651) 425-4058 or (800) 788-4085. Fax: (651) 452-9204.

Program Details

The *Read Naturally®* program can be implemented using one of four products: *Read Naturally® Masters Edition*, *Read Naturally® Encore*, *Read Naturally® Software Edition*, and *Read Naturally® Live*. These products share a common fluency-building strategy and are designed to supplement a school's core language arts instruction. The program aims to improve fluency, accuracy, and comprehension by increasing the time students spend reading and can be used during class time as a pull-out intervention during the school day or as part of an after-school program. The core strategy in all *Read Naturally®* products includes:

(I) Modeling of story reading. Students listen to, and read along with, a recording of a fluent reader reading a story to help students model correct pronunciation, rate, and expression.

(II) Repeated reading of text to develop oral reading fluency. Students engage in one-minute practice readings to build their mastery of the passage. Once students feel they can achieve their reading speed goal, they alert the teacher. The teacher then conducts a "pass timing" during which students are evaluated against four criteria: (1) student reaches goal rate, (2) student makes three or fewer errors, (3) passage is read with appropriate phrasing, and (4) comprehension questions are answered correctly. If students don't meet these criteria, they spend additional time practicing the reading of the passage, and then the teacher conducts the "pass timing" again.

(III) Progress monitoring. Students graph their scores to track their progress from the initial reading to the final reading of each story. The graphs also show students' progress over successive stories. These tools aim to ensure teacher and student awareness of each student's progress.

The four *Read Naturally®* products differ in (1) their delivery mode, (2) the specific sequenced texts used, and (3) whether phonics instruction is included. Read *Naturally® Masters Edition* and *Read Naturally® Encore* use audio CDs in conjunction with hard-copy reading materials. *Read Naturally® Software Edition* and *Read Naturally® Live* are computer- or web-based, respectively. The particular texts vary by product, but all include a series of sequenced texts. *Read Naturally® Software Edition*, *Read Naturally®Encore*, and *Read Naturally® Live* also include instruction in phonics.

Each *Read Naturally®* product includes a teacher's manual that includes the rationale for the program, descriptions of materials needed to implement the program, instructions for implementing the program, and lesson plans for introducing the program to students.

Cost

Individual *Read Naturally®* materials vary in price. Products using audio CDs (*Read Naturally® Masters Edition* or *Read Naturally® Encore*) cost $129 per set. *Read Naturally® Software Edition* costs $125 per reading level for one computer and $399 per level for a school network version. *Read Naturally® Live*, the online software version, is priced per seat, ranging from $149 for one seat to $1,999 for 130 seats. Teacher training is available at an additional cost. Additional materials, including timers, posters, glossaries, crossword puzzles, and assessment materials, are also available.

September 4, 2013

Tom Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121

Dear Mr. Ihnot,

The What Works Clearinghouse (WWC), an initiative of the U. S. Department of Education's Institute of Education Sciences, was established to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education.

The purpose of this letter is to notify you that we are in the process of revising the intervention reports that focus on Read Naturally® – Beginning Reading (update released July 2013), English Language Learners (update released July 2010), Students with Learning Disabilities (released July 2010), and Adolescent Literacy (released March 2013) – as discussed in recent communication from the WWC. In this letter, we ask you to review a brief intervention summary.

Intervention Summary. We have attached a brief summary of the features of Read Naturally®. We welcome your comment on its accuracy and invite you to provide additional information for the reports if appropriate.

We would appreciate receiving comments by mail (P.O. Box 2393, Princeton, NJ 08543-2393), email (info@whatworks.ed.gov) or fax (609-799-0005), by **September 18th** (10 business days from the date of this letter). Information received after this date may not be incorporated into the current revisions, but will be considered when the reports are updated.

Sincerely,

Neil Seftor
Co-Principal Investigator and Deputy Project Director
What Works Clearinghouse at Mathematica Policy Research
866-503-6114

Attachments: Intervention Summary

What Works

From:What Works
Sent:4 Sep 2013 15:07:16 -0400
To:Karla Ramy;info@readnaturally.com
Subject:WWC Description of Read Naturally
Attachments:Developer Contact Attachment_Read Naturally Description.pdf, Developer Contact Letter_Review of Read Naturally Program Description.pdf
Dear Mr. Ihnot,

The What Works Clearinghouse (WWC), an initiative of the U. S. Department of Education's Institute of Education Sciences, was established to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education. As such, we review studies on education interventions that may be included in our reports.

The purpose of the attached letter is to notify you that we are in the process of revising the four intervention reports that discuss Read Naturally® – Beginning Reading (update released July 2013), English Language Learners (update released July 2010), Students with Learning Disabilities (released July 2010), and Adolescent Literacy (released March 2013) – as discussed in recent communication from the WWC. In this letter, we ask you to review a brief intervention summary.

Sincerely,

Neil Seftor
Co-Principal Investigator and Deputy Project Director
What Works Clearinghouse at Mathematica Policy Research

School Psychology Review,
2010, Volume 39, No. 3, pp. 447–462

Curriculum-Based Measurement of Oral Reading: An Evaluation of Growth Rates and Seasonal Effects Among Students Served in General and Special Education

Theodore J. Christ
University of Minnesota

Benjamin Silberglitt
Technology and Information Educational Services

Seungsoo Yeo
Korean Educational Development Institute

Damien Cormier
University of Minnesota

Abstract. Curriculum-based measurement of oral reading (CBM-R) is often used to benchmark growth in the fall, winter, and spring. CBM-R is also used to set goals and monitor student progress between benchmarking occasions. The results of previous research establish an expectation that weekly growth on CBM-R tasks is consistently linear throughout the academic year. The patterns of CBM-R growth were examined for a large sample of students (N = 3808) from both general education and special education populations in second to sixth grades. Results support four general conclusions: (a) annual growth is more substantial within the general education population; (b) growth is more substantial in earlier elementary grades; (c) more growth occurs in the fall than the spring season (i.e., seasonal effect), especially within the early primary general education population; and (d) the seasonal effect is less pronounced within the special education population. Estimates of growth within and across seasons are presented and implications are discussed.

Curriculum-based measurement (CBM) is used to index annual student growth across the primary grades. Procedures and measurement metrics are developed for mathematics, spelling, written expression, and reading. CBM oral reading (CBM-R) rate is the most researched and well established of those available procedures (Wayman, Wallace, Wiley, Ticha, & Espin, 2007). CBM-R is used as an index of reading rate and fluency, which are identified as critical skills to target for instruction and intervention within the early stages of

Correspondence regarding this article should be addressed to Theodore J. Christ, 56 East River Road, Minneapolis, MN 55455; E-mail: tchrist@umn.edu

Copyright 2010 by the National Association of School Psychologists, ISSN 0279-6015

reading development (National Reading Panel, 2000; National Research Council, 1998). CBM-R outcome also functions as a robust indicator of overall reading development throughout the primary grades (Wayman et al., 2007).

CBM-R progress monitoring data are collected and used to evaluate instructional effects, determine when to modify the instruction, and evaluate response to intervention (Deno, 1985, 1986). Teachers who use progress monitoring data are likely to use more specific and measurable goals, rely more substantially on data to guide instruction, and modify instruction more frequently (Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, & Stecker, 1989). Progress monitoring applications of CBM-R have the potential to confer the most substantial benefit to students. Progress monitoring can be conceptualized as either interim benchmark/screening assessments when data are collected three to four times per year, or time series continuous assessments when data are collected daily or weekly.

Both interim and time series CBM-R data are plotted graphically and evaluated against goals and goal lines. Fuchs and Shinn (1989) recommended that the rate of expected growth should derive from local normative performance so that individual student achievement can be compared against the local normative performance of grade-level peers. A goal line is established by connecting the CBM-R data point for the observed level of performance from Week 1 (initial performance) to the expected/typical level of performance at Week 36 (ending performance). Once graphed, the goal line provides a graphic trajectory of expected growth for the academic year. Ongoing time series data are then plotted on the graph and evaluated against the goal line. For example, if a student's initial level of CBM-R performance in September was 20 words read correctly per minute (WRCM) and in May that student was expected to read 60 WRCM, then the expected annual growth is 40 WRCM across 8 months, or 32 weeks. That translates to ~1.25 WRCM per week of growth (1.25 WRCM = 40 WRCM/32 weeks). The goal line and trajectory of expected growth is based on the assumption of a monotonic and linear trend.

Monotonic Linear Growth

Standards of expected performance within and across the year provide the foundation to establish instructional goals and goal lines. There are two influential studies that established standards of weekly CBM-R growth (Deno, Fuchs, Marston, & Shin, 2001; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993), which have application in both research and practice. Fuchs et al. (1993) analyzed a CBM-R data set comprised of 103 students in general education and 14 students in special education. Each student was assessed weekly for an entire academic year using CBM-R procedures. Although some students evidenced a negatively accelerating annual growth, the researchers concluded that annual growth is, on average, monotonic and linear. They also observed a negatively accelerating trend across grades. That is, the rate of CBM-R growth was positive in all grades, but the rate of growth declined in each successive grade. The resulting analysis yielded grade-specific estimates of standard and ambitious rates of growth—and *assumed* a pattern of monotonic linear growth. Respectively, the growth, determined by the number of WRCM, were 1.5 and 2.0 in first and second grade, 1.0 and 1.5 in third, 0.85 and 1.1 in fourth, 0.5 and 0.8 in fifth, and 0.3 and 0.63 in sixth.

Deno et al. (2001) analyzed a large data set comprised of 2675 students in general education and 324 in special education classes. The student sample was from school districts across the country. The researchers relied on the *assumption* of monotonic linear annual growth. Ordinary least squares regression was used to calculate weekly growth. The study yielded useful estimates of observed and expected growth across grades and populations. The authors estimated that beginning readers in the general education population can be expected to improve 2 WRCM/week until they achieve 30 WRCM; thereafter, students in the general education population can be expect to improve at least 1 WRCM/week. Researchers observed that students in the special education

population demonstrated substantially less growth, which approximated 0.5–0.8 WRCM/week, unless they are provided with robust evidence-based instruction, which might improve the observed rate of growth to 1.39 WRCM/week.

Prior studies provide useful estimates of typical, ambitious, and expected rates of CBM-R growth; however, those estimates depend on the assumption that growth is monotonic and linear across the academic year. That assumption should be examined more closely to determine its veracity. The prior studies also rely on analyses that neglect the inherent nesting of students within classrooms and schools, which establish potential bias in the linear and quadratic functions derived.

Improved Modeling of Annual and Seasonal Growth

Research is necessary to evaluate alternate models of annual growth and test the predominant assumption that annual growth is monotonic and linear. This study will evaluate both linear (constant) and piece wise (nonconstant) models of growth from fall to winter to spring; moreover, this study will employ the superior approach of linear mixed model (LMM) rather than that of ordinary least squares used in prior research (Deno et al., 2001; Fuchs et al., 1993).

Ordinary least squares only estimates the fixed effects, which assumes that all cases have identical parameter estimates for intercept and growth. This is a limitation when modeling CBM-R growth because it is likely to be an erroneous assumption in most cases. The results of published research provide evidence for substantial magnitudes of variance in both intercept and growth among students in a sample (Shin, Deno, & Espin, 2000; Stage, 2001). It is likely that the intercept and growth rates should be treated as random effects in many instances. LMM can be used to estimate both the fixed effects and the random effects, which are the intercepts and growth rates for individual student cases in the sample. LMM takes into account the variances within and across individuals in the sample. Finally, ordinary least squares requires that missing data are removed casewise whereas LMM can handle missing data, which is likely to yield less biased estimates of intercept and slopes while enhancing power (Duncan, Duncan, & Strycker, 2006). This approach facilitates analysis of both seasonal effects, as pursued by two previous studies (Ardoin & Christ, 2008; Graney, Missall, Martinez, & Bergstrom, 2009), and compares those effects to that of linear growth as assumed in other studies (Deno et al., 2001; Fuchs et al., 1993).

The calendar year was divided into three seasons for the purpose of this study. The seasons were defined by those periods between standard tri-annual assessment occasions: fall, winter and spring. The duration between fall and winter assessments defines the fall season, that between winter and spring defines the spring season, and that between spring and fall defines the summer season. If growth is consistent and linear throughout the calendar year, then there is no *seasonal effect*. Conversely, if growth is not consistent and linear throughout the calendar year, then there is a seasonal effect whereby there is more growth observed for some seasons than for others.

The summer season is the most distinct for the majority of students because they are not exposed to formal instruction. The rate of achievement is likely to decline in the summer months (Kim, 2004) unless formal instruction is provided through summer school programs (Stage, 2001). A seasonal effect in the summer is expected because the conditions for learning are substantially different from that which occurs in the spring and fall seasons. The effect, however, does not seem specific to those differences between summer and the academic year. Ardoin and Christ (2008) observed statistically significant differences in CBM-R growth between the fall and spring seasons. That is, among a sample of second-grade students there was more growth in the fall than in the spring season. That result replicated across three combinations of CBM-R passages sampled from the Dynamic Indicators of Basic Early Literacy Skills. A subsequent study aimed to replicate and extend those findings

with a sample of third- through fifth-grade students across a variety of CBM curriculum domains (Graney et al., 2009). However, Graney et al. (2009) concluded that there was more CBM-R growth in the spring season than in the fall season, failing to replicate the work of Ardoin and Christ. Graney et al. speculated that the failure to replicate findings across studies might relate to the specification of data collection schedules, sample characteristics, issues of instrumentation, or the distribution of instructional intensity across the academic year. Although it is difficult to make a single attribution, it should be noted that the student sample used by Graney et al. derived from a school–university partnership program that included features to promote data use and evidence-based interventions. Indeed, that program might influence the pattern of student growth; however, additional research and replications are necessary to help determine whether specific patterns for seasonal effects should be expected.

There are limitations to the analytic procedures used in the two prior studies. Both generated estimates of growth from difference scores across seasons. That is, growth was estimated as the difference between performance at Time 1 and Time 2, which was then divided by the number of intervening weeks to yield an estimate of weekly growth. Although this method is intuitive, and it was useful for the preliminary work to examine seasonal effects, the method of difference scores to estimate change has long been criticized as an unreliable method of analysis (Cronbach & Furby, 1970) and may have contributed to the inconsistent findings. The analytic procedures used in this study improve on those of prior studies that examined seasonal effects and modeled annual growth. The analytic procedures used in this study will compare the hypothesized seasonal effect directly with that of linear growth.

Purpose

The purposes of this study were to (a) estimate the rate of growth within and across seasons for general education and special education populations; (b) evaluate whether the rate of CBM-R growth is likely to be consistent (i.e., linear model) or inconsistent (i.e., piecewise model) throughout the academic year; and (c) test the hypothesis that the seasonal effect is typically characterized by more robust growth in the fall than for the spring season. That hypothesis is counter to the assumptions inherent to the linear models as applied in many previous studies and it is counter to the findings of Graney et al. (2009).

Method

Participants and Setting

A total of 4824 students (52% male, 48% female) in the second through sixth grades were included in the sample. The ethnicity breakdown of the sample was 94% White, not Hispanic; 2% Native American; 2% Asian or Pacific Islander; 1% Hispanic; and 1% Black, not Hispanic. Participants were enrolled in seven elementary schools, located within five school districts in the rural and suburban Midwest. The sample consisted of 31% students receiving free or reduced-price lunch, and 8% students receiving special education services.

As part of an ongoing benchmark assessment plan and response to intervention implementation (see Bollman, Silberglitt, & Gibbons, 2007, for additional information), all students in participating schools were assessed using CBM-R every fall, winter, and spring of every year across the grade levels studied. All students for whom relevant data were collected participated in the study. Data were gathered between the years of 2001 and 2005 for each student in the sample. Data from across multiple years were available for many of the students in the sample; thus many students participated in the study at multiple grade levels.

Measures

The CBM-R measures were developed or selected to be curriculum-independent, grade-level appropriate, and of equivalent difficulty across probes within each grade level. Each probe contained approximately 250

words typed on an 8 by 10 piece of paper in a grade-appropriate font (range = 12–14 point with larger print for the second and third grades). Probes were selected from a standardized set of commercially available published passages (AIMSweb; Howe & Shinn, 2002). The technical manual for AIMSweb passages (Howe & Shinn, 2002) and peer-reviewed research (Wayman et al., 2007) provide evidence that test–retest reliability and alternate-form reliabilities approximate or exceed .90. Criterion-related validity coefficients typically approximate the .70–.90 range. The same set of CBM-R probes were used across all districts for all 4 years of the study.

Procedure

CBM-R data were collected in 4-month intervals in the fall, winter, and spring of each academic year. Standardized administration and scoring procedures were implemented for each administration, which was conducted by trained school-based personnel and high school honors students. These individuals were trained to administer three successive CBM-R probes and score WRCM using standardized procedures that have appeared in the published literature (Howe & Shinn, 2002; Shinn, 1988, 1989). Administration occurred outside of the classroom to minimize distractions, in a separate room or quiet hallway while the administrator and student were seated at desks or in chairs.

Training included a brief instructional session that lasted approximately 1 hr. Each training session was followed by an assessment of competency that included practice and evaluation of scoring procedures. No formal criterion for scoring was established until the final years of the study. In 2003–2004 and 2004–2005, each scorer was required to come within 2 WCPM of the correct score on three consecutive videotaped assessments. Although data were not maintained, few scorers required more than one opportunity to meet criterion. Finally, inter-rater reliability was not evaluated; however, CBM-R procedures have established high levels of inter-rater reliability that appear consistently within the published literature. A survey of 10 arbitrarily selected CBM-R studies that were published between 1995 and 2005 reported inter-rater reliabilities that were consistently above .95.

Analytical Plan

A LMM analysis was performed to investigate evidence of seasonal effects for CBM-R. The present LMM analysis considered two hypothesized growth models: linear and piecewise (Duncan et al., 2006; Flora, 2008). The linear model assumes a static growth rate across time points—meaning that regardless of change over time, the estimated growth rate from the linear model is identical to the sample's growth rate. In contrast, the piecewise model assumes that the developmental trajectory is not consistent across time, making it possible to provide two different linear trends. Thus, the piecewise model takes into account a possible nonlinear trajectory for CBM-R. Based on the three CBM-R data points used in this study, the piecewise model can be parameterized into two levels of analyses. First, Level 1 can be expressed as

$$y_{it} = \beta_{0i} + \beta_{1i}\lambda_{1t} + \beta_{2i}\lambda_{2t} + \varepsilon_{it} \quad (1)$$

where y_{it} is the outcome of CBM-R for subject i at time t, β_{0i} is the value of CBM-R for subject i in fall, β_{1i} is the first linear trend of subject i from fall to winter, λ_{1t} is the value of the linear trend at time t (Week 0 and Week 18), β_{2i} is the second linear trend of subject i from winter to spring, λ_{2t} is the value of the linear trend at time t (Week 18 and Week 36), and ε_{it} is unique measurement error for subject i at time t.

Level 2 can be described as

$$\beta_{0i} = \beta_0\zeta_{0i} \quad (2)$$

$$\beta_{1i} = \beta_1\zeta_{1i} \quad (3)$$

$$\beta_{2i} = \beta_2\zeta_{2i} \quad (4)$$

where β_0 represents the average initial status for all subjects, β_1 refers to the average of the first linear growth rate for all subjects, and β_2

refers to the average of the second slope trend for all subjects. ζ_{0i}, ζ_{1i}, and ζ_{2i} refer to the random deviation of the initial status, the first growth rate, and the second growth rate, respectively. As a result, the piecewise model presented earlier assumes that the second slope is statistically different from the first slope, indicating that there is a seasonal effect for CBM-R throughout the academic year.

The first step in this analysis was to compare the two models (linear and piecewise models) using fit indices, which included the likelihood ratio test (LRT), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The three fit indices are considered to be "goodness-of-fit" indices, where the smallest fit statistics are viewed as the best model. If results from these goodness-of-fit indices reveal that the linear model is better than the piecewise model, then the slope is the same from fall to spring and there is no evidence of a significant seasonal effect for CBM-R across the academic year. Conversely, if goodness-of-fit indices are better for the piecewise model than the linear model, then we conclude that the slopes are different from fall to winter and from winter to spring and there is evidence of a seasonal effect across the academic year. If the piecewise model is the better fit, then an additional model is needed to estimate the extent to which the first slope differs from the second slope.

The additional model, which is called *added growth model,* would be identical to the piecewise model described earlier, except for the supplementary time factor loading matrix (Flora, 2008). For example, in the piecewise model, the time factor loading can be described as

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 18 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (5)$$

where the first column is set at 1, meaning that the outcome of the fall CBM-R score (intercept) does not change across time. The second column refers to the first linear time metric indicating change in CBM-R score from fall to winter (Flora, 2008). Similarly, the third column represents change in CBM-R score from winter to fall.

For the added growth model, the time factor loading can be expressed as:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 36 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (6)$$

The matrix in Equation 6 is identical to the matrix in Equation 5, with the exception of the second column with 0, 18, and 36. The different time factor loadings for the additional model is applied to examine the extent to which linear change in CBM-R from fall to winter increases or decreases compared to that from winter to spring. In other words, the added growth model analysis makes it possible to contrast the first linear slope and the second linear slope. The results from the additional analysis can also be viewed as an effect size estimate between the first and second linear slope (Flora, 2008).

Data Preparation

Overall, the rates of missing data for fall, winter, and spring seasons were less than 4% for all grade levels. All LMM analyses were conducted using the SAS MIXED program with full maximum likelihood estimation that automatically conducts all adjustments for missing data. Monthly CBM data used in this study were nested within districts. Therefore, a three-level LMM was used to satisfy an assumption that data collected in this study might not be independent among school districts. An alpha level for all statistical tests in this study was set at .001.

Model assumptions were evaluated, which included a test of normality using Kolmogorov-Smirnove statistics, skew, and kurtosis, and homoscedasticity of residuals by plotting raw residuals at each level (Levels 1 and 2). No violations were detected. The Level 3 LMM analysis was used to test whether CBM data used in the study were nested within school districts. As a result, variations in the slopes and intercepts between school districts were not significantly different

Table 1
Descriptive Statistics for the Level and Change of CBM-R

	Seasonal Benchmark Levels			Weekly Change	
Population	Fall *M* (*SD*)	Winter *M* (*SD*)	Spring *M* (*SD*)	Fall Season (Fall–Winter) *M* (*SD*)	Spring Season (Winter–Spring) *M* (*SD*)
General education					
Second	57 (32)	88 (34)	107 (35)	1.71 (.73)	1.02 (.64)
Third	82 (37)	107 (38)	125 (40)	1.38 (.79)	0.97 (.75)
Fourth	106 (38)	128 (39)	141 (41)	1.19 (.72)	0.78 (.70)
Fifth	125 (40)	145 (41)	160 (42)	1.12 (.76)	0.83 (.75)
Sixth	142 (41)	158 (42)	171 (43)	0.89 (.71)	0.74 (.75)
Total	99 (47)	122 (45)	137 (46)	1.26 (.74)	0.87 (.72)
Special education					
Second	33 (28)	55 (35)	74 (38)	1.17 (.81)	1.08 (.63)
Third	52 (33)	72 (39)	88 (41)	1.09 (.81)	0.89 (.78)
Fourth	77 (40)	96 (43)	108 (45)	0.99 (.72)	0.67 (.65)
Fifth	89 (43)	106 (45)	118 (47)	0.90 (.76)	0.67 (.79)
Sixth	106 (47)	119 (50)	132 (51)	0.70 (.72)	0.73 (.69)
Total	72 (45)	89 (48)	104 (49)	0.97 (.77)	0.81 (.71)

Note. CBM-R = curriculum-based measurement of oral reading. Skew and kurtosis were within reasonable limits ($z \leq 2.00$) for all conditions. *N* values for second, third, fourth, fifth, and sixth grades were, respectively, 1790, 1885, 1913, 1942, and 1144 for general education; 111, 134, 170, 192, and 116 for special education.

from zero for all grade levels, meaning that both the slopes and intercepts at the district levels were identical. In addition, such a result supports evidence that the Level 2 LMM analysis introduced in the method section is appropriate for this study.

Results

Summaries of the means and standard deviations for CBM-R at each time point are presented in Table 1. As shown in Table 1, the average growth as measured with CBM-R was positive for all conditions across seasons (fall–winter, winter–spring), grade levels (second to sixth) and populations (general education, special education).

Comparison Between Linear Model and Piecewise Model

Results for the comparisons of linear and piecewise models are presented in Table 2. For students in general education, the piecewise model provided the best fit for tri-annual assessment data, as evidenced by the magnitudes of LRT, AIC, and BIC observed for the linear and piecewise models respectively. The χ^2 difference test between the two models also revealed that the piecewise model fit significantly better than the linear model.

For students in special education, on the basis of the LRT criterion, the piecewise model for all grade levels was viewed as the best-fitting model. Significant differences in LRTs between the two models, however, were found in the second through fourth grades. There were inconsistencies with respect to AIC and BIC criteria, which might indicate the linear model fit better in some cases within the sample of special education students. Although the results of AIC and BIC are noted, the LRT was the primary test/criterion used here.

Table 2
Deviation Between Linear and Piecewise Models for General Education Students and Special Education Students

	Linear Model			Piecewise Model			Differences (Linear-Piecewise Model)		
Grade	LRT	AIC	BIC	LRT	AIC	BIC	LRT	AIC	BIC
General education									
Second	46231.7	46239.7	46261.7	45542.6	45556.6	45595.0	689.1*	683.1	666.7
Third	49566.9	49574.9	49597.0	49357.7	49371.7	49410.5	209.2*	203.2	186.5
Fourth	49768.1	49776.1	49798.3	49520.1	49534.1	49573.0	248.0*	242.0	225.3
Fifth	50974.9	50982.9	51005.2	50866.6	50880.6	50919.6	108.3*	102.3	85.6
Sixth	29892.9	29900.9	29921.1	29873.6	29887.6	29922.9	19.3*	13.3	−1.8
Special education									
Second	2816.3	2824.3	2835.1	2802.6	2816.6	2835.6	13.7*	7.7	−0.5
Third	3501.8	3507.8	3516.5	3487.9	3495.9	3507.5	13.9*	11.9	9.0
Fourth	4394.9	4402.9	4415.4	4378.9	4392.9	4414.9	16.0*	10.0	0.5
Fifth	5090.6	5098.6	5111.6	5083.3	5097.3	5120.1	7.3	1.3	−8.5
Sixth	3048.2	3056.2	3067.2	3046.3	3060.3	3079.5	1.9	−4.1	−12.3

Note. LRT = likelihood ratio test; AIC = Akaike information criterion; BIC = Bayesian information criterion.
*$p < .001$.

Comparison Between the First Slope and Second Slope for General and Special Education

Table 3 presents estimates of both the first (fall–winter) and second (winter–spring) linear slope. In the cases of both students in general education and special education, all of the seasonal linear slopes were positive with magnitudes significantly different than zero ($p < .001$). That observation was consistent with the conclusion that reading achievement accelerated continuously throughout the academic year for all student samples. Although the mean rate of growth across seasons was consistently greater for students in general education as compared with students in special education, it was the first slopes that were most distinct. Further, the differences in the second slopes varied by grade level.

Additional analysis was conducted to examine whether the first linear slopes (fall) were significantly different from the second slopes (spring). Table 3 presents the differences between the two linear slopes along with the results of significance tests. Figure 1 depicts the slopes and model differences.

In the sample of general education students, statistically significant differences ($p < .001$) were observed between the first and second slopes in all grade-level conditions (Table 3). That is, the rate of growth for all grade levels dropped in the second part of the academic year. It is relevant to note that the magnitude of difference between the two slopes (seasonal effects) and standardized effect sizes consistently decreased as grade level increased. This might indicate that the seasonal effect in reading is larger in the lower primary grades than the upper primary grades. The power estimates for detecting seasonal effects for students in general education were very large for all grade levels (range = 0.96–1.00). This indicates a good chance of detecting a false null hypothesis (linear model) when the alternative model (piecewise model) is correct (Duncan et al., 2006).

In the sample of special education students, the average of the first slope (fall) was

Table 3
Comparison Between the First Slope and the Second Slope for General Education Students and Special Education Students

Grade	Intercept $\hat{\beta}_0$ (*SE*)	Fall/First Slope (Fall to Winter) $\hat{\beta}_1$ (*SE*)	Spring/ Second Slope (Winter to Spring) $\hat{\beta}_2$ (*SE*)	Difference Between Fall Slope and Spring Slopes $\hat{\beta}_2{}^a$ (*SE*)	Standardized Effect Sizes[b]	Power
General education						
Second	57.42* (.76)	1.71* (.02)	1.02* (.01)	0.69* (.02)	34.50	1.00
Third	82.31* (.85)	1.38* (.02)	0.97* (.02)	0.41* (.03)	13.60	1.00
Fourth	105.78* (.93)	1.19* (.02)	0.78* (.02)	0.41* (.12)	3.41	1.00
Fifth	124.43* (.89)	1.12* (.02)	0.83* (.02)	0.29* (.12)	2.41	1.00
Sixth	142.49* (1.21)	0.88* (.02)	0.74* (.02)	0.14* (.16)	0.87	0.96
Special education						
Second	34.02* (2.67)	1.17* (.08)	1.08* (.06)	0.09 (.10)	0.90	0.86
Third	50.85* (2.81)	1.09* (.07)	0.89* (.06)	0.20 (.10)	2.00	0.87
Fourth	75.50* (2.99)	0.99* (.05)	0.67* (.05)	0.32* (.08)	4.00	0.92
Fifth	86.28* (3.06)	0.90* (.05)	0.67* (.06)	0.23 (.09)	2.50	0.54
Sixth	106.98* (4.40)	0.69* (.07)	0.73* (.06)	−0.04 (.10)	0.40	0.11

Note. SE = standard error.
[a]The estimated coefficients of the second slope from an added growth model.
[b]Standardized effect sizes = $\hat{\beta}_2$/SE, which is a strategy developed by Satorra and Saris (1985), was used to calculate power.
*$p < .001$.

greater than the second slope (spring) in all grades except for sixth; however, those differences were not statistically significant except in the fourth-grade condition (Table 3). Moreover, the pattern of reduced seasonal effects for progressively higher grades did not generalize to the sample of special education students. Instead, the magnitude was greatest in fourth grade (difference in slopes = 0.32, $p < .001$), somewhat smaller in third (0.20, $p > .001$) and fifth (0.23, $p > .001$), and hardly perceptible in second (0.09, $p > .001$) and sixth grades (−0.04, $p < .001$). The range of standardized effect sizes and power was variable across grades. Based on the results, it appears that, in general, the observed slopes in the sample of special education students were relatively consistent across seasons—in all conditions but fourth grade.

Discussion

Many of the published estimates of CBM-R annual growth rates rely on the assumption that the pattern is consistent and linear throughout the academic year (Deno et al., 2001; Fuchs et al., 1993). Although there is evidence for seasonal effects within some studies (Ardoin & Christ, 2009; Fuchs et al., 1993; Graney et al., 2009), the phenomenon of seasonal effects are generally ignored within the professional literature. This study yielded (a) estimates for the magnitude CBM-R annual growth along with those within the fall and spring seasons, and estimates were derived for both general education and special education populations; (b) both linear and piecewise growth models; and (c) hypothesis tests to examine whether fall growth is typically more substantial than spring growth.

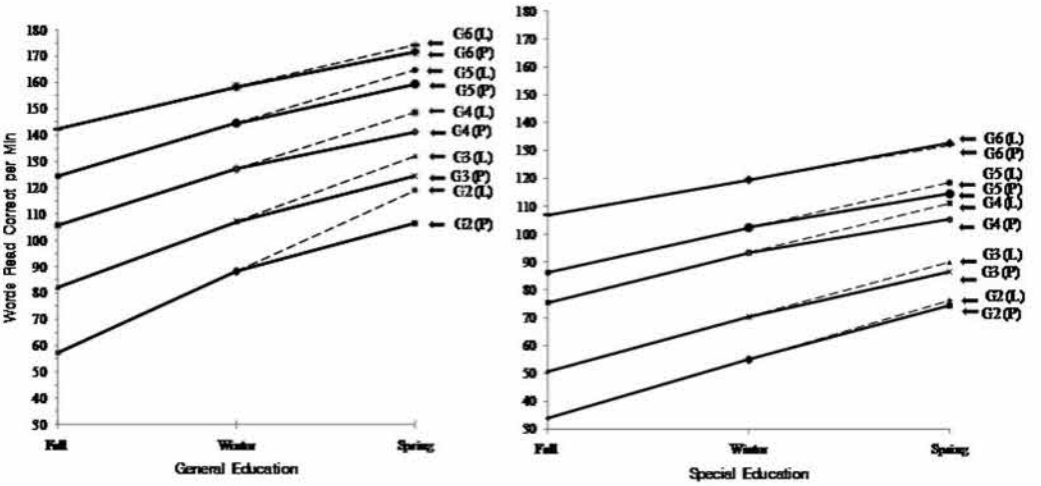


Figure 1. Differences for linear (L, solid line) and piecewise (P, dashed line) models of growth (G) for students in second through sixth grade within either the general (left) or special education population (right)

CBM-R growth was modeled and analyzed for five primary grades (second through sixth) and two student populations within and across seasons.

The results of this study support four general conclusions. Descriptive data make clear that, first, annual growth is greater among students in the general education population than for those in the special education population, especially in the fall season; and, second, more growth occurs in the early grades with less in the upper grades. Modeling and inferential analysis provide support for the third and fourth conclusions that follow. That is (third conclusion), there is a significant *seasonal effect* such that more growth occurs in the fall season than in the spring season for general education, but the magnitude of that effect declines with each progressive grade level. Finally, the seasonal effect is less pronounced among students in the special education population. In general, linear growth should not be assumed, seasonal effects should be expected—at least within the general education population—and piecewise models tend to fit better than linear models of growth.

Grade and Population Effects

Larger magnitudes of growth were observed for students in general education as compared to special education. This outcome was observed across five districts where CBM-R was used to screen, monitor, and evaluate instructional effects as part of a problem solving model. The systematic and ongoing use of CBM-R for these purposes did not establish equivalent growth rates between the student populations. Despite inconsistent growth rates across student groups, the rate of growth among students in the special education population was greater than observed in other studies (Deno et al., 2001; Fuchs et al., 1993). For example, the most recent estimates for average growth among students in the special education population were ~0.60 WRCM/week (Deno et al., 2001). That rate of growth was relatively consistent across grades. In contrast, the average growth estimates among students in the sample of students in special education used in this study were within the range of 0.67–1.17 WRCM/week (Table 3). Although cross-study comparisons cannot be used to establish that differences in systems-level practices (e.g., problem solving) and instructional practices were causally related to

the observed differences in growth rates, future research should investigate the phenomena more closely.

The results of this study converge with previous research (Deno et al., 2001; Fuchs et al., 1993) to support the conclusion that there is more CBM-R growth within the early primary grades. This phenomenon was consistent for both fall and spring seasons in both the general education and special education populations (Table 1). There was a steady decline for both general education and special education in weekly growth as grade level was increased. There was a corresponding decline in the magnitude of seasonal effects as grade levels increased. Results also support the conclusion that seasonal effects are more robust for students served within general education and less robust for students served within special education.

Seasonal Effect

The primary purpose of this research was to determine whether the rate of CBM-R growth is consistent across the year. Notwithstanding limitations, a seasonal effect was sufficiently robust and consistent across grades and student populations to warrant serious consideration. Estimates of weekly CBM-R growth within season are reported in Table 1. They are also depicted in Figure 1. Both visual analysis and statistical analysis of growth models support the conclusion that there was a *difference in the average rate of growth across seasons*, and that difference occurred in all grades, but was generally more robust in the lower grades of general education (Figure 1). With few exceptions, more growth occurred in the fall than in the spring, which is depicted as the differences in the slopes within Table 3.

The *magnitudes of seasonal effects were inconsistent* across grades and student populations. In general education, the difference between fall and spring slopes ranged from 0.14 to 0.69 per week, and all the differences were statistically significant with substantial effect sizes (Table 3). That effect is illustrated in Figure 1, which reveals that the interactions between seasonal effect and student population were most distinct in second grade. That is, the seasonal effect was most robust for the sample of students in second-grade general education classes, with a difference in growth from fall to winter of 0.69, and minor in comparison for students in second-grade special education classes, with a difference of 0.09.

The results of this study support the conclusion that growth differs between fall and spring seasons, which is inconsistent with a number of currently held views in the field. The consistency of the results across grades for general education begs the question of why this effect may be occur and what are the potential implications. There are a number of alternate hypotheses that offer some insight on the phenomenon and directions for future research.

Explanatory Hypotheses

This study was designed to examine *whether* seasonal effects are likely to occur in cases of CBM-R tri-annual assessment and challenge the assumption of linear growth. Indeed, the results do provide support for seasonal effects; however, this study was not designed to examine the cause of seasonal effects. That is left to future research to isolate the cause(s) and potential remedies to accelerate reading achievement within seasons with deficit growth. The result of both this study and Ardoin and Christ (2008) converge to provide evidence that the seasonal effect is likely to coincide with deficit levels of growth in the latter part of the academic year, or the spring season, especially for the students in the early primary general education population. It seems that changes in instructional conditions, or student motivation, might explain deficit rates of growth within the spring season unless there is an active intervention to prevent those changes. It seems that data-based decision making could function to reduce, or reverse, the pattern of seasonal effects that were replicated across this study and that of Ardoin and Christ.

Teacher expectations. There are a number of ways in which teacher expectations can have an influence on student behavior,

which might be conveyed through discrepant use of praise and punishment statements, content-relevant feedback, individualized attention, differentiated instruction, and variability in the opportunities to respond and engage in instructionally relevant practice (Good, 1981). Prior research yielded conclusions that there are indeed sometimes differences in teacher treatment of high- and low-achieving students over the course of the year, which affected the amount and type of instruction and feedback that students received at different points in the academic year (Good et al., 1980). Researchers concluded that teachers are more invested in influencing and shaping student behavior at the beginning of the school year than later in the year. Furthermore, the differentiation in teacher behavior may have sustained effects and become more influential as the school year progresses, resulting in lower achievement and growth later in the school year in low-achieving students. Other research also demonstrated the potential influence of teacher expectations, including expectation biases across ethnic groups (McKown & Weinstein, 2007) and ability groups (Rosenthal & Jacobson, 1966).

Classroom management. There is evidence that classroom management and organizational strategies are implemented with greater fidelity and effort in the early part of the academic year (Cameron, Connor, & Morrison, 2004). The same research provides evidence that that those educators who approach classroom management and organization with the greatest fervor early in the year are those same educators whose practices drop off most substantially later in the academic year. Improved levels of engagement and instructional gain are clearly associated with good classroom management and organization, and are among better predictors of student achievement (Brophy, 1987). Taken together, research provides a foundation to postulate that differential levels of classroom management and organization across seasons might influence student achievement. This effect is hypothesized to be the greatest at lower grade levels, as younger students require greater direction in organization practices (Cameron et al., 2004). The pattern of seasonal effects observed in this study coincides with a classroom management and organizational hypothesis.

Intervention to accelerate seasonal growth. The results of this study indicate that the seasonal effect with deficit growth in the spring season was most pronounced in the middle primary grades (second and third) and were minimized for the special education population. The results of Graney et al. (2009) indicate that the seasonal effect can function so that there is deficit growth in the fall season and accelerated growth in the spring. As noted in the Introduction, the data set that was analyzed by Graney et al. (2009) derived from a university and school partnership to implement a data-based decision making and evidence-based interventions. Those factors might have influenced the patterns of seasonal effects. There is a substantial literature base to support the conclusion that data-based decision making can function to accelerate academic growth (Carnine & Granzin, 2001; Fuchs & Fuchs, 1986; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Fuchs, Fuchs, Hamlett, & Whinnery, 1991; Graney et al., 2009), and there is substantial theoretical support for the use of evidence-based practices, which were both components of service delivery that were targeted by the partnership program. It might be that progress monitoring and data use accelerated learning and ameliorated the deficit growth associated with seasonal effects. Indeed, students who are served within special education have explicit goals and their progress is monitored. Although it is not possible to draw any final conclusions, it may well be that the lack of convergence in seasonal patterns between Graney et al. and other studies—along with minimal seasonal effects observed for the special education population—provides further evidence for the use of data-based decision making to improve growth during otherwise deficit periods.

Limitations and Future Research

The sample studied was entirely from five relatively ethnically homogeneous dis-

tricts in a single midwestern state. Future research should attempt to corroborate these findings across a more diverse sample of students and educational programs. Finally, the results of this study derive from data collected at three points in the academic year (fall, winter, spring). Future research should examine whether the same conclusions would result from the analysis of a more robust data set (e.g., weekly, monthly) and with more precise dates that define the data collection schedule. Graney et al. (2009) correctly criticize the methods of this study and that of Ardoin and Christ (2008) because the precise dates of data collection were not used in the analysis; rather, both studies rely on data that are centered by month and not day. Although the authors of this and the prior study believe the conclusions would be substantially similar with different centering methods or more precise dates, the criticism ought to be accounted for in future research. Future research should investigate the nature of growth across summative and formative measurement conditions. Research should also examine the effect on educational decisions when the potential for seasonal effects is ignored and consistent annual growth is assumed.

Implications for School-Based Decision Making

CBM-R was developed to guide routine instructional decisions and enhance instructional effects (Deno, 1986), and it has emerged as one of the primary methods of assessment to guide problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention models of service delivery (National Association of State Directors of Education, 2005). As described in the Introduction of this study, CBM-R is used to establish expectations for both the level and rate of reading achievement throughout the primary grades; therefore, the results of this study provide at least two clear implications. First, there is good reason to question the validity of linear growth estimates as criteria to guide progress monitoring (i.e., discrepant in rate) and dual-discrepancy decisions (i.e., discrepant in both level and rate; Fuchs, Fuchs, McMaster & Al Otaiba, 2003). Second, seasonal effects can be either admired as a phenomenon or targeted for intervention. These implications are addressed in what follows.

First, if CBM-R estimates of growth are used to evaluate instructional effects and response to intervention, then those who interpret data should be aware that the magnitude of growth might be less in the spring than in the fall. There are two weighty implications that derive from this study. First, a dual-discrepancy approach to special education eligibility, which compares the observed level and rate of achievement to criterion standards (Fuchs, Fuchs, McMaster & Al Otaiba, 2003), might function to overidentify students during deficit seasons of growth and underidentify students during accelerated seasons of growth. If the data follow the pattern observed in this study, then it is likely that the dual-discrepancy model might underidentify students in the fall season and overidentify students in the spring season. That consequence is counter to the values of early intervention and prevention inherent within problem solving and response to intervention. Researchers and practitioners should pay close attention to the patterns of growth and the potential for seasonal effects to influence decisions at the local level. Accelerated rates of growth should be expected in the fall.

Second, the pattern of seasonal effects that are observed across populations and studies provide preliminary evidence that goal setting, progress monitoring and data-based decision making can minimize, eliminate or reverse the pattern of seasonal effects (e.g., Graney et al., 2009). Although future research is necessary, it is a working thesis that the systematic use of data to evaluate instructional effects and select effective instructional strategies effectively accelerates learning and prevents problems at the individual, group, and systems level. After all, data-based decision making is a central tenet of problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention (National Association of State Directors of Education, 2005); however, data must be both systematically collected *and* sys-

tematically used to establish effects (Fuchs et al., 1992, 1991; Fuchs & Fuchs, 1986). The mere act of data collection without data use typically fails to establish effects (Fuchs & Fuchs, 1986). The explanatory hypotheses proposed in the previous section establish that deficit growth in the spring season might be remediated at the systems level if data are systematically collected and systematically used. If detected, this problem (seasonal effect) is ripe for problem analysis at the local level. Deficit growth within a single season should not be admired as an acceptable phenomenon, but, instead, it should be targeted for analysis and intervention.

Christ (2008) defined problem analysis as the "collection, summary, and use of information to systematically test, reject, or verify relevant hypotheses to establish problem solutions" (p. 159). That is, deficit rates of growth in the spring season might be remediated if expectations remain high and data are used to evaluate and maintain high standards for instructional effects. There is no reason to assume that the seasonal effect is inherent to the educational system. Instead, hypotheses related to the causal and maintaining variables should be developed—such as those presented earlier—and tested through intervention to remediate the problem. It is likely that the pattern of seasonal effects is causally related to what occurs in the classroom; and what occurs in the classroom can be influenced by the systematic used of data to evaluate of response to intervention for universal, supplemental, and intensive tiers of service delivery. This study provides evidence of seasonal effects as a potential target for intervention at the system, grade, or classroom level.

References

AIMSweb. (2002). *Reading curriculum-based measurement (R-CBM) standard oral reading fluency assessment passages.* Retrieved October 2003, from http://www.aimsweb.com

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum based measurement slope estimate using data from tri-annual universal screenings *School Psychology Review, 37,* 109–125.

Bollman, K. A., Silberglitt, B., & Gibbons, K. A. (2007). The SCRED Model: Incorporating systems-level organization and a multi-tiered problem-solving process for intervention delivery. In M. Burns, S. Jimerson, & A. VanDerHayden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 321–332). New York: Springer.

Brophy, J. (1987). Synthesis of research on strategies for motivating students to learn. *Educational Leadership, 45*(2), 40–48.

Cameron, C. E., Connor, C. M., & Morrison, F. J. (2004). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology, 43,* 61–85.

Carnine, D., & Granzin, A. (2001). Setting learning expectations for students with disabilities. *School Psychology Review, 30,* 466–472.

Christ, T. J. (2008). Best practices in problem analysis. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 159–176). Bethesda, MD: National Association of School Psychologists.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74,* 68.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219–232.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15,* 358–374.

Deno, S. L. (2002). Problem solving as best practices. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 37–56). Bethesda, MD: National Association of School Psychologists.

Deno, S. L. (2005). Problem solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 10–42). New York: Guilford Press.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30,* 507–524.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). An introduction to latent variable growth curve modeling: Concepts, issues, and applications (2nd ed.). Mahwah, NY: Lawrence Erlbaum Associates.

Flora, D. B. (2008). Specifying piecewise latent trajectory models for longitudinal data. *Structural Equation Modeling, 15,* 513–533.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53,* 3–14.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10,* 43–52.

Fuchs, L. S., Fuchs, D., & Stecker, P. M. (1989). Effects of curriculum-based measurement on teachers' instructional planning. *Journal of Learning Disabilities, 22,* 51–59.

Fuchs, L. S., & Shinn, M. R. (1989). Writing CBM IEP objectives. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 131–152). New York: Guildford Press.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58,* 436–450.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22,* 27–48.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Whinnery, K. (1991). Effects of goal line feedback on level, slope, and stability of performance within curriculum-based measurement. *Learning Disabilities Research and Practice, 6,* 66–74.

Fuchs, D., Fuchs, L. S., McMaster, K. N., & Al Otaiba, S. (2003). Identifying children at risk for reading failure: Curriculum-based measurement and the dual-discrepancy approach. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 431–449). New York: Guilford Press.

Good, T. L. (1981). Teacher expectations and student perceptions: A decade of research. *Educational Leadership, 38*(5), 415–422.

Good, T., Cooper, H., & Blakey, X. (1980). Classroom interaction as a function of teacher expectations, student sex and time of year. *Journal of Educational Psychology, 72*, 378–385.

Graney, S. B., Missall, K. N., Martinez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School* Psychology, 47, 121–142.

Howe, K. B., & Shinn, M. M. (2002). Standard *reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Retrieved October 2004, from http://www.aimsweb.com

Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education for Students Placed at Risk, 9,* 169–188.

McKown, C., Weinstein, R. S. (2007). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*, 235–261.

National Association of State Directors of Education. (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: Author.

National Reading Panel. (2000). *Report of the National Reading Panel Subgroups: Teaching children to read* (No. NIH 00–4754). Bethesda, MD: National Institute of Child Health and Human Development.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports, 19,* 115–118.

Satorra, A., & Saris, W. E. (1985). The power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50*, 83–90.

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education, 34,* 164–172.

Shinn, M. R. (1988). Development of curriculum-based local norms for use in special education decision-making. *School Psychology Review, 17,* 61–80.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children.* New York: Guildford Press.

Shinn, M. (2008). Best practices for using curriculum based measurement in a problem solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 243–262). Bethesda, MD: National Association of School Psychologists.

Stage, S. A. (2001). Program evaluation using hierarchical linear modeling with curriculum-based measurement reading probes. *School Psychology Quarterly, 16,* 91–112.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41,* 85–120.

Date Received: May 30, 2007
Date Accepted: April 21, 2010
Action Editor: Tanya Eckert ■

Theodore J. Christ, PhD, is an associate professor in the Educational Psychology Department at the University of Minnesota in Minneapolis—Saint Paul. He is the codirector of the Research Institute for Problem Solving. He is engaged in research and training to evaluate, develop, and enhance the tools and procedures that enable data-based decision making and problem solving within school settings. Interests include progress monitoring, curriculum-based measurement, direct behavior ratings, and computer adaptive testing.

Benjamin Silberglitt completed his PhD in Educational Psychology at the University of Minnesota at Minneapolis—Saint Paul. His areas of interest include student assessment school organization and system-level change, problem solving, response to intervention, and data-based decision making. He is director of Software Applications for TIES, a technology cooperative of 39 Minnesota school districts.

Seungsoo Yeo, PhD, is a researcher at Korean Educational Development Institute. He received his PhD from the University of Minnesota at Twin Cities in 2008. His primary research interests include curriculum-based measurement, response to intervention, and learning disabilities. His methodological work has focused on latent growth modeling and multilevel meta-analysis.

Damien C. Cormier, MA, is a school psychology doctoral student in the Educational Psychology Department at the University of Minnesota. He is a graduate research assistant at the National Center on Educational Outcomes. His research interests are in the areas of large-scale assessment policy and practice, measurement issues in individual assessment, high-functioning autism, and cultural and linguistic biases in assessment.

Copyright of School Psychology Review is the property of National Association of School Psychologists and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Karla Ramy

From:Karla Ramy
Sent:5 Sep 2013 15:14:25 +0000
To:What Works
Subject:To Neil Seftor
Attachments:Christ Silberglitt Yeo Cormier (2010).pdf, Heistad.docx, EditedDeveloper Contact Attachment_Read Naturally Description.docx

Neil Seftor,

I have made a couple of minor changes in the program description for Read Naturally. (see attached)

In an August 26, 2013 letter from What Works Clearinghouse (WWC), Jill Constantine said:

> *Order of findings.* We appreciate and agree with your suggestion of July 17, in which you propose that the WWC list findings in order of their strength, consistent with the presentation on the web. Therefore, we will revise the order of findings listed in the text on the front page of the report, the associated table, and the website summary for the three reports with findings in more than one domain. We anticipate that the revised intervention reports will be posted to the web by the end of September. (Reference :QR2010013, WWC2304)

I look forward to this change being executed in September, 2013.

Please email a verification that the change in order of findings will read:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

Also, I sent an email to Jill Constantine on August 26, 2013 requesting WWC to change its Beginning Reading effectiveness rating on reading fluency for Read Naturally, Inc. from mixed effects to positive effects. (See below)

Will my request in the August 26, 2013 email be reviewed now so that the Read Naturally revisions will include an effectiveness rating change from mixed effects to positive effects for reading fluency?

I look forward to the changes. It is critical that educators get accurate and up to date information.
Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Jill Constantine,

I am requesting that What Works Clearinghouse (WWC) change its Beginning Reading effectiveness rating on reading fluency for Read Naturally, Inc. from mixed effects to positive effects for several reasons. Please consider and respond to the following information.

First, in a recent WWC review of a study by Dr. David Heistad, statistically significant gains in oral reading fluency for third grade students across multiple schools were not considered in the evaluation of the effects of a Read Naturally intervention on fluency. After reviewing the most recent WWC Procedures and Standards Handbook (February, 2013), the reason for exclusion may be due to a simple misunderstanding.

Page 15 of the WWC Procedures and Standards Handbook (Outcome Eligibility and Reporting section) states the following:

"....outcome measures is that they not be over aligned with the intervention. When outcome measures are closely aligned with or tailored to the intervention, the study findings may not be an accurate indication of the effect of the intervention. For example, an outcome measure based on an assessment that relied on materials used in the intervention condition but not in the comparison condition...."

The measure of oral reading fluency in Dr. Heistad's study was *The Reading Fluency Monitor*. It should be noted that *The Reading Fluency Monitor* is not a component of the Read Naturally Intervention. The *Reading Fluency Monitor* was developed by RMC Research Corporation, located in Portsmouth, New Hampshire. Dr. Heistad's implication that *The Reading Fluency Monitor* was developed by Read Naturally, Inc. was a miscommunication in the Test Instruments section of his study. Heistad's verification of this misunderstanding is attached.

Second, the Arvans study summary analysis underestimated the effect of the Read Naturally group on oral reading fluency. The Read Naturally group had a large effect size of .81 for fluency. The control EAU group had a moderate effect size of .57 for fluency. Both groups had excellent fluency gains. It should be noted that the use of the Read Naturally Intervention was extensive at the school before the study was conducted. The teachers were well versed in the Read Naturally strategy. Students used Read Naturally in the first and second grade in the two years prior to the study. Also, some students in the EAU group used Read Naturally during the study. That is, it seems as though the EAU group may not have been demonstrative of a true control group. The difference between the Read Naturally group's .81 effect size and EAU's .57 effect size was .24 after eight weeks. An effect size difference of .24 is significant when extrapolated over a school year.

Ethan R. Van Norman, M.A., performed a similar analysis to what appeared in the Christ and Davie study (2008). In that study, the authors first calculated a slope estimate from three time points for each

student in the EAU and the experimental group. The slope estimate represented the number of words read correct per minute (WCPM) improvement per week. The mean and standard deviation of slopes were then calculated for each group. The percent of improvement of the experimental group in relation to the EAU group was calculated. After this, the authors used the percent of improvement and applied it to an aggressive rate of growth (1.50 WCPM improvement per week). That value was then multiplied by 36, and 1.50 was multiplied by 36. The difference between those values was then interpreted as a hypothetical effect if the Read Naturally Intervention was delivered across an entire school year.

A similar analysis was conducted on the Arvans dataset. Slope estimates were calculated for each student from two observations eight weeks apart. The mean slope value for the experimental group was 2.92 WCPM improvement per week (SD = 1.54) compared to the EAU group, which had a mean slope estimate of 2.24 (SD = 2.36). The .68 difference in mean slope for the Read Naturally Intervention group represents a 30% improvement over the EAU group. Assuming an aggressive rate of growth of 1.50 WCPM for typical students, a 30% increase would translate to a 1.95 rate of growth. Extended across 36 weeks, this represents a net increase of 70 WCPM compared to 54 WCPM with a typical rate of growth (1.50 WCPM improvement per week across 36 weeks).

Although not ideal, slope estimates from two time points have been used to summarize growth in previous CBM-R research studies (e.g., Christ, Silberglit, Yeo & Cormier, 2010). See the attached document for further details.

Another way to analyze the Arvans study is to use the Hasbrouck-Tindal Oral Reading Fluency Norms (http://www.readnaturally.com/howto/orftable.htm). The table suggests that third grade students at the 50th percentile have an average weekly improvement of 1.1 WCPM. The performance of the Read Naturally group in the Arvans study is also impressive when compared to the typical rate of fluency growth of third grade students.

Ethan R. Van Norman, M.A., did an analysis to extrapolate growth in the same manner using normative values. For third grade students, weekly growth estimates for students in the 50th percentile typically approximate 1.10 WCPM per week. Assuming the Read Naturally group's 30% improvement over the EAU group and the intervention is delivered for 36 weeks, a student in the 50th percentile, on average would be improving at a rate of 1.43 WCPM per week. After 36 weeks this would translate to an improvement of 51 WCPM compared to 40 WCPM if the student did not receive the intervention. This is a substantial difference.

Considering all of this information, Read Naturally programs should receive a positive effects rating for fluency. A timely reevaluation is critical so that WWC may change Read Naturally's effectiveness rating for fluency from mixed effects to positive effects. Please respond to the points I've raised in this email and let me know if you need further information.

Sincerely,

Tom Ihnot
CEO

Intervention Summary for Read Naturally®

Please review the intervention description below, and tell us whether we have described the intervention accurately, and whether we have included all relevant details. The information you provide may be incorporated into reports on Read Naturally® and posted to the WWC website for public access.

Program Description

The Read Naturally® program is a supplemental reading program that aims to improve the reading fluency, accuracy, and comprehension of elementary, middle school, high school and adult students using a combination of texts, audio CDs, and computer software. The program uses one of four products that share a common fluency-building strategy: Read Naturally® Masters Edition, Read Naturally® Encore, Read Naturally® Software Edition, and Read Naturally® Live. The common strategy includes: modeling of story reading, repeated reading of text for developing oral reading fluency, and systematic monitoring of student progress by teachers and the students themselves. Students work at their own reading level, progress through the program at their own rate, and work (for the most part) on an independent basis. The program can be delivered in three ways: (1) students use audio CDs with hard-copy reading materials (Read Naturally® Masters Edition, Read Naturally® Encore), (2) students use the computer-based version (Read Naturally® Software Edition), or (3) students use the web-based version (Read Naturally® Live). This intervention report includes studies of Read Naturally® Masters Edition and Read Naturally®
Software Edition.

Background

Developed by Candyce Ihnot, the four Read Naturally® products are distributed by Read Naturally, Inc. Address: 2945 Lone Oak Drive, Suite #190, Saint Paul, MN 55121. Email: info@readnaturally.com. Web: www.readnaturally.com. Telephone: (651) 425-4058 or (800) 788-4085. Fax: (651) 452-9204.

Program Details

The Read Naturally® program can be implemented using one of four products: Read Naturally® Masters Edition, Read Naturally® Encore, Read Naturally® Software Edition, and Read Naturally® Live. These products share a common fluency-building strategy and are designed to supplement a school's core language arts instruction. The program aims to improve fluency, accuracy, and comprehension by increasing the time students spend reading and can be used during class time, as a pull-out intervention during the school day or as part of an after-school program. The core strategy in all Read Naturally® products includes:

(I) Modeling of story reading. Students listen to, and read along with, a recording of a fluent reader reading a story to help students model correct pronunciation, rate, and expression.

(II) Repeated reading of text to develop oral reading fluency. Students engage in one-minute practice readings to build their mastery of the passage. Once students feel they can achieve their reading speed goal, they alert the teacher. The teacher then conducts a "pass timing" during which students are evaluated against four criteria: (1) student reaches goal rate, (2) student makes three or fewer errors, (3) passage is read with appropriate phrasing, and (4) comprehension questions are answered correctly. If students don't meet these criteria, they spend additional time practicing the reading of the passage,

and then the teacher conducts the "pass timing" again.

(III) Progress monitoring. Students graph their scores to track their progress from the initial reading to the final reading of each story. The graphs also show students' progress over successive stories. These tools aim to ensure teacher and student awareness of each student's progress.

The four Read Naturally® products differ in (1) their delivery mode, (2) the specific sequenced texts used, and (3) whether phonics instruction is included. Read Naturally® Masters Edition and Read Naturally® Encore use audio CDs in conjunction with hard-copy reading materials. Read Naturally® Software Edition and Read Naturally® Live are computer- or web-based, respectively. The particular texts vary by product, but all include a series of sequenced texts. Read Naturally® Software Edition, Read Naturally®Encore, and Read Naturally® Live also include instruction in phonics.

Each Read Naturally® product includes a teacher's manual that includes the rationale for the program, descriptions of materials needed to implement the program, instructions for implementing the program, and lesson plans for introducing the program to students.

Cost

Individual Read Naturally
®
materials vary in price. Products using audio CDs (Read Naturally
®
Masters
Edition or Read Naturally
®
Encore) cost $129 per set. Read Naturally
®
Software Edition costs $125 per
reading level for one computer and $399 per level for a school network version. Read Naturally
®
Live, the
online software version, is priced per seat, ranging from $149 for one seat to $1,999 for 130 seats. Teacher training is available at an additional cost. Additional materials, including timers, posters, glossaries, crossword puzzles, and assessment materials, are also available.

I was mistaken when I indicated that the Reading Fluency Monitor was developed by Read Naturally. It was actually developed by the RMC Research Corporation out of Portsmouth, New Hampshire.

Sorry for the misunderstanding,

- Dave

David Heistad, Ph.D., Executive Director
Research and Evaluation Department
Bloomington Public Schools - ISD 271
1350 W. 106th Street
Bloomington, MN 55431
ph: 952-681-6485
cell: 952-250-5971

What Works

From:What Works
Sent:6 Sep 2013 14:04:28 -0400
To:Karla Ramy
Subject:RE: To Neil Seftor
Dear Mr. Ihnot,

Thank you so much for your prompt response to our email and for submitting your comments on the program description of Read Naturally®. We are also modifying our presentation of findings as part of the revisions to the four intervention reports on Read Naturally®. We will let you know once the revised reports are available on the WWC website.

The WWC Quality Review Team is currently investigating the issue you raise regarding the effectiveness rating on reading fluency in the report on Read Naturally® released under the Beginning Reading topic area. We will communicate with you separately about that concern.

Sincerely,

Neil Seftor
Co-Principal Investigator and Deputy Project Director
What Works Clearinghouse at Mathematica Policy Research

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Thursday, September 05, 2013 11:14 AM
To: What Works
Subject: To Neil Seftor

Neil Seftor,

I have made a couple of minor changes in the program description for Read Naturally. (see attached)

In an August 26, 2013 letter from What Works Clearinghouse (WWC), Jill Constantine said:

> *Order of findings.* We appreciate and agree with your suggestion of July 17, in which you propose that the WWC list findings in order of their strength, consistent with the presentation on the web. Therefore, we will revise the order of findings listed in the text on the front page of the report, the associated table, and the website summary for the three reports with findings in more than one domain. We anticipate that the revised intervention reports will be posted to the web by the end of September. (Reference :QR2010013, WWC2304)

I look forward to this change being executed in September, 2013.

Please email a verification that the change in order of findings will read:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

Also, I sent an email to Jill Constantine on August 26, 2013 requesting WWC to change its Beginning Reading effectiveness rating on reading fluency for Read Naturally, Inc. from mixed effects to positive effects. (See below)

Will my request in the August 26, 2013 email be reviewed now so that the Read Naturally revisions will include an effectiveness rating change from mixed effects to positive effects for reading fluency?

I look forward to the changes. It is critical that educators get accurate and up to date information.
Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Jill Constantine,

I am requesting that What Works Clearinghouse (WWC) change its Beginning Reading effectiveness rating on reading fluency for Read Naturally, Inc. from mixed effects to positive effects for several reasons. Please consider and respond to the following information.

First, in a recent WWC review of a study by Dr. David Heistad, statistically significant gains in oral reading fluency for third grade students across multiple schools were not considered in the evaluation of the effects of a Read Naturally intervention on fluency. After reviewing the most recent WWC Procedures and Standards Handbook (February, 2013), the reason for exclusion may be due to a simple misunderstanding.

Page 15 of the WWC Procedures and Standards Handbook (Outcome Eligibility and Reporting section) states the following:

"....outcome measures is that they not be over aligned with the intervention. When outcome measures are closely aligned with or tailored to the intervention, the study findings may not be an accurate

indication of the effect of the intervention. For example, an outcome measure based on an assessment that relied on materials used in the intervention condition but not in the comparison condition…."

The measure of oral reading fluency in Dr. Heistad's study was *The Reading Fluency Monitor*. It should be noted that *The Reading Fluency Monitor* is not a component of the Read Naturally Intervention. The *Reading Fluency Monitor* was developed by RMC Research Corporation, located in Portsmouth, New Hampshire. Dr. Heistad's implication that *The Reading Fluency Monitor* was developed by Read Naturally, Inc. was a miscommunication in the Test Instruments section of his study. Heistad's verification of this misunderstanding is attached.

Second, the Arvans study summary analysis underestimated the effect of the Read Naturally group on oral reading fluency. The Read Naturally group had a large effect size of .81 for fluency. The control EAU group had a moderate effect size of .57 for fluency. Both groups had excellent fluency gains. It should be noted that the use of the Read Naturally Intervention was extensive at the school before the study was conducted. The teachers were well versed in the Read Naturally strategy. Students used Read Naturally in the first and second grade in the two years prior to the study. Also, some students in the EAU group used Read Naturally during the study. That is, it seems as though the EAU group may not have been demonstrative of a true control group. The difference between the Read Naturally group's .81 effect size and EAU's .57 effect size was .24 after eight weeks. An effect size difference of .24 is significant when extrapolated over a school year.

Ethan R. Van Norman, M.A., performed a similar analysis to what appeared in the Christ and Davie study (2008). In that study, the authors first calculated a slope estimate from three time points for each student in the EAU and the experimental group. The slope estimate represented the number of words read correct per minute (WCPM) improvement per week. The mean and standard deviation of slopes were then calculated for each group. The percent of improvement of the experimental group in relation to the EAU group was calculated. After this, the authors used the percent of improvement and applied it to an aggressive rate of growth (1.50 WCPM improvement per week). That value was then multiplied by 36, and 1.50 was multiplied by 36. The difference between those values was then interpreted as a hypothetical effect if the Read Naturally Intervention was delivered across an entire school year.

A similar analysis was conducted on the Arvans dataset. Slope estimates were calculated for each student from two observations eight weeks apart. The mean slope value for the experimental group was 2.92 WCPM improvement per week (SD = 1.54) compared to the EAU group, which had a mean slope estimate of 2.24 (SD = 2.36). The .68 difference in mean slope for the Read Naturally Intervention group represents a 30% improvement over the EAU group. Assuming an aggressive rate of growth of 1.50 WCPM for typical students, a 30% increase would translate to a 1.95 rate of growth. Extended across 36 weeks, this represents a net increase of 70 WCPM compared to 54 WCPM with a typical rate of growth (1.50 WCPM improvement per week across 36 weeks).

Although not ideal, slope estimates from two time points have been used to summarize growth in previous CBM-R research studies (e.g., Christ, Silberglit, Yeo & Cormier, 2010). See the attached document for further details.

Another way to analyze the Arvans study is to use the Hasbrouck-Tindal Oral Reading Fluency Norms (http://www.readnaturally.com/howto/orftable.htm). The table suggests that third grade students at the 50th percentile have an average weekly improvement of 1.1 WCPM. The performance of the Read Naturally group in the Arvans study is also impressive when compared to the typical rate of fluency growth of third grade students.

Ethan R. Van Norman, M.A., did an analysis to extrapolate growth in the same manner using normative values. For third grade students, weekly growth estimates for students in the 50th percentile typically approximate 1.10 WCPM per week. Assuming the Read Naturally group's 30% improvement over the EAU group and the intervention is delivered for 36 weeks, a student in the 50th percentile, on average would be improving at a rate of 1.43 WCPM per week. After 36 weeks this would translate to an improvement of 51 WCPM compared to 40 WCPM if the student did not receive the intervention. This is a substantial difference.

Considering all of this information, Read Naturally programs should receive a positive effects rating for fluency. A timely reevaluation is critical so that WWC may change Read Naturally's effectiveness rating for fluency from mixed effects to positive effects. Please respond to the points I've raised in this email and let me know if you need further information.

Sincerely,

Tom Ihnot
CEO

October 18, 2013

Mr. Thomas Ihnot
CEO
Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
kramy@readnaturally.com
info@readnaturally.com

Reference: QR2013002

Dear Mr. Ihnot:

Thank you for your email to Dr. Constantine regarding the What Works Clearinghouse rating of effectiveness on the reading fluency domain for Read Naturally®. In response to your email, we conducted an independent quality review to address the concerns you have raised. The WWC quality review team responds to concerns raised about WWC reviews published on our website. When a quality review is conducted, a researcher who was not involved in the initial review undertakes an independent assessment of the studies in question. The researcher also investigates the procedures used and decisions made during the original review of the studies. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld on every review conducted by the WWC.

The quality review investigated two concerns raised in your email: the classification of the oral reading fluency measure in Heistad's (2008) study as overaligned with the intervention and the effect size on the oral reading fluency measure in the Arvans (2010) study.

The quality review's investigation of the review of the Heistad (2008) study concluded that the WWC followed its procedures in determining whether the fluency measure was overaligned. The decision about whether a measure is overaligned is made by the WWC's lead methodologist for the topic area in consultation with the topic area's lead content expert (*WWC Procedures and Standards Handbook*, v3.0, page 15). In this case, they decided that the measure was overaligned with the intervention, as words used in the Read Naturally® curriculum also appear in the Reading Fluency Monitor.

The quality review's investigation of the review of the Arvans (2010) study concluded that the WWC followed its procedures in calculating the effect size for the fluency measure. The WWC's default approach is to compute the effect size numerator as the difference between the pre- and posttest mean difference for the intervention group and the pre- and posttest mean difference for the comparison group (*WWC Procedures and Standards Handbook*, v3.0, page F.4). In this case, the difference-in-differences adjustment reduced the final treatment-control difference from 7.17 words to 4.73 words, due to the initial 2.44 word difference in favor of the treatment group. The

effect size computed on the impact of 4.73 words was correctly reported as 0.16 standard deviations. With regard to presenting the effects of a full year of the intervention based on the eight-week posttest, the WWC does not conduct or report extrapolations, but does clearly indicate the timeframe for outcome measurement in the report.

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

Sincerely,

(b)(6)

Jill Constantine
Director, What Works Clearinghouse

cc: Joy Lesnick, IES

What Works

From:What Works
Sent:18 Oct 2013 15:54:17 +0000
To:'kramy@readnaturally.com';'info@readnaturally.com'
Cc:'Joy.Lesnick@ed.gov'
Subject:What Works Clearinghouse (2013002)
Attachments:QRT2013002 Response_10.18.2013.pdf
Dear Mr. Ihnot:

Attached is a response to the questions you raised on August 26, 2013 in your inquiries concerning the What Works Clearinghouse Beginning Reading effectiveness rating on reading fluency for Read Naturally®.

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/.

School Psychology Review,
2010, Volume 39, No. 3, pp. 447–462

Curriculum-Based Measurement of Oral Reading: An Evaluation of Growth Rates and Seasonal Effects Among Students Served in General and Special Education

Theodore J. Christ
University of Minnesota

Benjamin Silberglitt
Technology and Information Educational Services

Seungsoo Yeo
Korean Educational Development Institute

Damien Cormier
University of Minnesota

Abstract. Curriculum-based measurement of oral reading (CBM-R) is often used to benchmark growth in the fall, winter, and spring. CBM-R is also used to set goals and monitor student progress between benchmarking occasions. The results of previous research establish an expectation that weekly growth on CBM-R tasks is consistently linear throughout the academic year. The patterns of CBM-R growth were examined for a large sample of students (N = 3808) from both general education and special education populations in second to sixth grades. Results support four general conclusions: (a) annual growth is more substantial within the general education population; (b) growth is more substantial in earlier elementary grades; (c) more growth occurs in the fall than the spring season (i.e., seasonal effect), especially within the early primary general education population; and (d) the seasonal effect is less pronounced within the special education population. Estimates of growth within and across seasons are presented and implications are discussed.

Curriculum-based measurement (CBM) is used to index annual student growth across the primary grades. Procedures and measurement metrics are developed for mathematics, spelling, written expression, and reading. CBM oral reading (CBM-R) rate is the most researched and well established of those available procedures (Wayman, Wallace, Wiley, Ticha, & Espin, 2007). CBM-R is used as an index of reading rate and fluency, which are identified as critical skills to target for instruction and intervention within the early stages of

Correspondence regarding this article should be addressed to Theodore J. Christ, 56 East River Road, Minneapolis, MN 55455; E-mail: tchrist@umn.edu

Copyright 2010 by the National Association of School Psychologists, ISSN 0279-6015

reading development (National Reading Panel, 2000; National Research Council, 1998). CBM-R outcome also functions as a robust indicator of overall reading development throughout the primary grades (Wayman et al., 2007).

CBM-R progress monitoring data are collected and used to evaluate instructional effects, determine when to modify the instruction, and evaluate response to intervention (Deno, 1985, 1986). Teachers who use progress monitoring data are likely to use more specific and measurable goals, rely more substantially on data to guide instruction, and modify instruction more frequently (Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, & Stecker, 1989). Progress monitoring applications of CBM-R have the potential to confer the most substantial benefit to students. Progress monitoring can be conceptualized as either interim benchmark/screening assessments when data are collected three to four times per year, or time series continuous assessments when data are collected daily or weekly.

Both interim and time series CBM-R data are plotted graphically and evaluated against goals and goal lines. Fuchs and Shinn (1989) recommended that the rate of expected growth should derive from local normative performance so that individual student achievement can be compared against the local normative performance of grade-level peers. A goal line is established by connecting the CBM-R data point for the observed level of performance from Week 1 (initial performance) to the expected/typical level of performance at Week 36 (ending performance). Once graphed, the goal line provides a graphic trajectory of expected growth for the academic year. Ongoing time series data are then plotted on the graph and evaluated against the goal line. For example, if a student's initial level of CBM-R performance in September was 20 words read correctly per minute (WRCM) and in May that student was expected to read 60 WRCM, then the expected annual growth is 40 WRCM across 8 months, or 32 weeks. That translates to ~1.25 WRCM per week of growth (1.25 WRCM = 40 WRCM/32 weeks). The goal line and trajectory of expected growth is based on the assumption of a monotonic and linear trend.

Monotonic Linear Growth

Standards of expected performance within and across the year provide the foundation to establish instructional goals and goal lines. There are two influential studies that established standards of weekly CBM-R growth (Deno, Fuchs, Marston, & Shin, 2001; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993), which have application in both research and practice. Fuchs et al. (1993) analyzed a CBM-R data set comprised of 103 students in general education and 14 students in special education. Each student was assessed weekly for an entire academic year using CBM-R procedures. Although some students evidenced a negatively accelerating annual growth, the researchers concluded that annual growth is, on average, monotonic and linear. They also observed a negatively accelerating trend across grades. That is, the rate of CBM-R growth was positive in all grades, but the rate of growth declined in each successive grade. The resulting analysis yielded grade-specific estimates of standard and ambitious rates of growth—and *assumed* a pattern of monotonic linear growth. Respectively, the growth, determined by the number of WRCM, were 1.5 and 2.0 in first and second grade, 1.0 and 1.5 in third, 0.85 and 1.1 in fourth, 0.5 and 0.8 in fifth, and 0.3 and 0.63 in sixth.

Deno et al. (2001) analyzed a large data set comprised of 2675 students in general education and 324 in special education classes. The student sample was from school districts across the country. The researchers relied on the *assumption* of monotonic linear annual growth. Ordinary least squares regression was used to calculate weekly growth. The study yielded useful estimates of observed and expected growth across grades and populations. The authors estimated that beginning readers in the general education population can be expected to improve 2 WRCM/week until they achieve 30 WRCM; thereafter, students in the general education population can be expect to improve at least 1 WRCM/week. Researchers observed that students in the special education

population demonstrated substantially less growth, which approximated 0.5–0.8 WRCM/week, unless they are provided with robust evidence-based instruction, which might improve the observed rate of growth to 1.39 WRCM/week.

Prior studies provide useful estimates of typical, ambitious, and expected rates of CBM-R growth; however, those estimates depend on the assumption that growth is monotonic and linear across the academic year. That assumption should be examined more closely to determine its veracity. The prior studies also rely on analyses that neglect the inherent nesting of students within classrooms and schools, which establish potential bias in the linear and quadratic functions derived.

Improved Modeling of Annual and Seasonal Growth

Research is necessary to evaluate alternate models of annual growth and test the predominant assumption that annual growth is monotonic and linear. This study will evaluate both linear (constant) and piece wise (nonconstant) models of growth from fall to winter to spring; moreover, this study will employ the superior approach of linear mixed model (LMM) rather than that of ordinary least squares used in prior research (Deno et al., 2001; Fuchs et al., 1993).

Ordinary least squares only estimates the fixed effects, which assumes that all cases have identical parameter estimates for intercept and growth. This is a limitation when modeling CBM-R growth because it is likely to be an erroneous assumption in most cases. The results of published research provide evidence for substantial magnitudes of variance in both intercept and growth among students in a sample (Shin, Deno, & Espin, 2000; Stage, 2001). It is likely that the intercept and growth rates should be treated as random effects in many instances. LMM can be used to estimate both the fixed effects and the random effects, which are the intercepts and growth rates for individual student cases in the sample. LMM takes into account the variances within and across individuals in the sample. Finally, ordinary least squares requires that missing data are removed casewise whereas LMM can handle missing data, which is likely to yield less biased estimates of intercept and slopes while enhancing power (Duncan, Duncan, & Strycker, 2006). This approach facilitates analysis of both seasonal effects, as pursued by two previous studies (Ardoin & Christ, 2008; Graney, Missall, Martinez, & Bergstrom, 2009), and compares those effects to that of linear growth as assumed in other studies (Deno et al., 2001; Fuchs et al., 1993).

The calendar year was divided into three seasons for the purpose of this study. The seasons were defined by those periods between standard tri-annual assessment occasions: fall, winter and spring. The duration between fall and winter assessments defines the fall season, that between winter and spring defines the spring season, and that between spring and fall defines the summer season. If growth is consistent and linear throughout the calendar year, then there is no *seasonal effect*. Conversely, if growth is not consistent and linear throughout the calendar year, then there is a seasonal effect whereby there is more growth observed for some seasons than for others.

The summer season is the most distinct for the majority of students because they are not exposed to formal instruction. The rate of achievement is likely to decline in the summer months (Kim, 2004) unless formal instruction is provided through summer school programs (Stage, 2001). A seasonal effect in the summer is expected because the conditions for learning are substantially different from that which occurs in the spring and fall seasons. The effect, however, does not seem specific to those differences between summer and the academic year. Ardoin and Christ (2008) observed statistically significant differences in CBM-R growth between the fall and spring seasons. That is, among a sample of second-grade students there was more growth in the fall than in the spring season. That result replicated across three combinations of CBM-R passages sampled from the Dynamic Indicators of Basic Early Literacy Skills. A subsequent study aimed to replicate and extend those findings

with a sample of third- through fifth-grade students across a variety of CBM curriculum domains (Graney et al., 2009). However, Graney et al. (2009) concluded that there was more CBM-R growth in the spring season than in the fall season, failing to replicate the work of Ardoin and Christ. Graney et al. speculated that the failure to replicate findings across studies might relate to the specification of data collection schedules, sample characteristics, issues of instrumentation, or the distribution of instructional intensity across the academic year. Although it is difficult to make a single attribution, it should be noted that the student sample used by Graney et al. derived from a school–university partnership program that included features to promote data use and evidence-based interventions. Indeed, that program might influence the pattern of student growth; however, additional research and replications are necessary to help determine whether specific patterns for seasonal effects should be expected.

There are limitations to the analytic procedures used in the two prior studies. Both generated estimates of growth from difference scores across seasons. That is, growth was estimated as the difference between performance at Time 1 and Time 2, which was then divided by the number of intervening weeks to yield an estimate of weekly growth. Although this method is intuitive, and it was useful for the preliminary work to examine seasonal effects, the method of difference scores to estimate change has long been criticized as an unreliable method of analysis (Cronbach & Furby, 1970) and may have contributed to the inconsistent findings. The analytic procedures used in this study improve on those of prior studies that examined seasonal effects and modeled annual growth. The analytic procedures used in this study will compare the hypothesized seasonal effect directly with that of linear growth.

Purpose

The purposes of this study were to (a) estimate the rate of growth within and across seasons for general education and special education populations; (b) evaluate whether the rate of CBM-R growth is likely to be consistent (i.e., linear model) or inconsistent (i.e., piecewise model) throughout the academic year; and (c) test the hypothesis that the seasonal effect is typically characterized by more robust growth in the fall than for the spring season. That hypothesis is counter to the assumptions inherent to the linear models as applied in many previous studies and it is counter to the findings of Graney et al. (2009).

Method

Participants and Setting

A total of 4824 students (52% male, 48% female) in the second through sixth grades were included in the sample. The ethnicity breakdown of the sample was 94% White, not Hispanic; 2% Native American; 2% Asian or Pacific Islander; 1% Hispanic; and 1% Black, not Hispanic. Participants were enrolled in seven elementary schools, located within five school districts in the rural and suburban Midwest. The sample consisted of 31% students receiving free or reduced-price lunch, and 8% students receiving special education services.

As part of an ongoing benchmark assessment plan and response to intervention implementation (see Bollman, Silberglitt, & Gibbons, 2007, for additional information), all students in participating schools were assessed using CBM-R every fall, winter, and spring of every year across the grade levels studied. All students for whom relevant data were collected participated in the study. Data were gathered between the years of 2001 and 2005 for each student in the sample. Data from across multiple years were available for many of the students in the sample; thus many students participated in the study at multiple grade levels.

Measures

The CBM-R measures were developed or selected to be curriculum-independent, grade-level appropriate, and of equivalent difficulty across probes within each grade level. Each probe contained approximately 250

words typed on an 8 by 10 piece of paper in a grade-appropriate font (range = 12–14 point with larger print for the second and third grades). Probes were selected from a standardized set of commercially available published passages (AIMSweb; Howe & Shinn, 2002). The technical manual for AIMSweb passages (Howe & Shinn, 2002) and peer-reviewed research (Wayman et al., 2007) provide evidence that test–retest reliability and alternate-form reliabilities approximate or exceed .90. Criterion-related validity coefficients typically approximate the .70–.90 range. The same set of CBM-R probes were used across all districts for all 4 years of the study.

Procedure

CBM-R data were collected in 4-month intervals in the fall, winter, and spring of each academic year. Standardized administration and scoring procedures were implemented for each administration, which was conducted by trained school-based personnel and high school honors students. These individuals were trained to administer three successive CBM-R probes and score WRCM using standardized procedures that have appeared in the published literature (Howe & Shinn, 2002; Shinn, 1988, 1989). Administration occurred outside of the classroom to minimize distractions, in a separate room or quiet hallway while the administrator and student were seated at desks or in chairs.

Training included a brief instructional session that lasted approximately 1 hr. Each training session was followed by an assessment of competency that included practice and evaluation of scoring procedures. No formal criterion for scoring was established until the final years of the study. In 2003–2004 and 2004–2005, each scorer was required to come within 2 WCPM of the correct score on three consecutive videotaped assessments. Although data were not maintained, few scorers required more than one opportunity to meet criterion. Finally, inter-rater reliability was not evaluated; however, CBM-R procedures have established high levels of inter-rater reliability that appear consistently within the published literature. A survey of 10 arbitrarily selected CBM-R studies that were published between 1995 and 2005 reported inter-rater reliabilities that were consistently above .95.

Analytical Plan

A LMM analysis was performed to investigate evidence of seasonal effects for CBM-R. The present LMM analysis considered two hypothesized growth models: linear and piecewise (Duncan et al., 2006; Flora, 2008). The linear model assumes a static growth rate across time points—meaning that regardless of change over time, the estimated growth rate from the linear model is identical to the sample's growth rate. In contrast, the piecewise model assumes that the developmental trajectory is not consistent across time, making it possible to provide two different linear trends. Thus, the piecewise model takes into account a possible nonlinear trajectory for CBM-R. Based on the three CBM-R data points used in this study, the piecewise model can be parameterized into two levels of analyses. First, Level 1 can be expressed as

$$y_{it} = \beta_{0i} + \beta_{1i}\lambda_{1t} + \beta_{2i}\lambda_{2t} + \varepsilon_{it} \quad (1)$$

where y_{it} is the outcome of CBM-R for subject i at time t, β_{0i} is the value of CBM-R for subject i in fall, β_{1i} is the first linear trend of subject i from fall to winter, λ_{1t} is the value of the linear trend at time t (Week 0 and Week 18), β_{2i} is the second linear trend of subject i from winter to spring, λ_{2t} is the value of the linear trend at time t (Week 18 and Week 36), and ε_{it} is unique measurement error for subject i at time t.

Level 2 can be described as

$$\beta_{0i} = \beta_0\zeta_{0i} \quad (2)$$

$$\beta_{1i} = \beta_1\zeta_{1i} \quad (3)$$

$$\beta_{2i} = \beta_2\zeta_{2i} \quad (4)$$

where β_0 represents the average initial status for all subjects, β_1 refers to the average of the first linear growth rate for all subjects, and β_2

refers to the average of the second slope trend for all subjects. ζ_{0i}, ζ_{1i}, and ζ_{2i} refer to the random deviation of the initial status, the first growth rate, and the second growth rate, respectively. As a result, the piecewise model presented earlier assumes that the second slope is statistically different from the first slope, indicating that there is a seasonal effect for CBM-R throughout the academic year.

The first step in this analysis was to compare the two models (linear and piecewise models) using fit indices, which included the likelihood ratio test (LRT), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The three fit indices are considered to be "goodness-of-fit" indices, where the smallest fit statistics are viewed as the best model. If results from these goodness-of-fit indices reveal that the linear model is better than the piecewise model, then the slope is the same from fall to spring and there is no evidence of a significant seasonal effect for CBM-R across the academic year. Conversely, if goodness-of-fit indices are better for the piecewise model than the linear model, then we conclude that the slopes are different from fall to winter and from winter to spring and there is evidence of a seasonal effect across the academic year. If the piecewise model is the better fit, then an additional model is needed to estimate the extent to which the first slope differs from the second slope.

The additional model, which is called *added growth model,* would be identical to the piecewise model described earlier, except for the supplementary time factor loading matrix (Flora, 2008). For example, in the piecewise model, the time factor loading can be described as

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 18 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (5)$$

where the first column is set at 1, meaning that the outcome of the fall CBM-R score (intercept) does not change across time. The second column refers to the first linear time metric indicating change in CBM-R score from fall to winter (Flora, 2008). Similarly, the third column represents change in CBM-R score from winter to fall.

For the added growth model, the time factor loading can be expressed as:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 36 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (6)$$

The matrix in Equation 6 is identical to the matrix in Equation 5, with the exception of the second column with 0, 18, and 36. The different time factor loadings for the additional model is applied to examine the extent to which linear change in CBM-R from fall to winter increases or decreases compared to that from winter to spring. In other words, the added growth model analysis makes it possible to contrast the first linear slope and the second linear slope. The results from the additional analysis can also be viewed as an effect size estimate between the first and second linear slope (Flora, 2008).

Data Preparation

Overall, the rates of missing data for fall, winter, and spring seasons were less than 4% for all grade levels. All LMM analyses were conducted using the SAS MIXED program with full maximum likelihood estimation that automatically conducts all adjustments for missing data. Monthly CBM data used in this study were nested within districts. Therefore, a three-level LMM was used to satisfy an assumption that data collected in this study might not be independent among school districts. An alpha level for all statistical tests in this study was set at .001.

Model assumptions were evaluated, which included a test of normality using Kolmogorov-Smirnove statistics, skew, and kurtosis, and homoscedasticity of residuals by plotting raw residuals at each level (Levels 1 and 2). No violations were detected. The Level 3 LMM analysis was used to test whether CBM data used in the study were nested within school districts. As a result, variations in the slopes and intercepts between school districts were not significantly different

Table 1
Descriptive Statistics for the Level and Change of CBM-R

	Seasonal Benchmark Levels			Weekly Change	
Population	Fall *M* (*SD*)	Winter *M* (*SD*)	Spring *M* (*SD*)	Fall Season (Fall–Winter) *M* (*SD*)	Spring Season (Winter–Spring) *M* (*SD*)
General education					
Second	57 (32)	88 (34)	107 (35)	1.71 (.73)	1.02 (.64)
Third	82 (37)	107 (38)	125 (40)	1.38 (.79)	0.97 (.75)
Fourth	106 (38)	128 (39)	141 (41)	1.19 (.72)	0.78 (.70)
Fifth	125 (40)	145 (41)	160 (42)	1.12 (.76)	0.83 (.75)
Sixth	142 (41)	158 (42)	171 (43)	0.89 (.71)	0.74 (.75)
Total	99 (47)	122 (45)	137 (46)	1.26 (.74)	0.87 (.72)
Special education					
Second	33 (28)	55 (35)	74 (38)	1.17 (.81)	1.08 (.63)
Third	52 (33)	72 (39)	88 (41)	1.09 (.81)	0.89 (.78)
Fourth	77 (40)	96 (43)	108 (45)	0.99 (.72)	0.67 (.65)
Fifth	89 (43)	106 (45)	118 (47)	0.90 (.76)	0.67 (.79)
Sixth	106 (47)	119 (50)	132 (51)	0.70 (.72)	0.73 (.69)
Total	72 (45)	89 (48)	104 (49)	0.97 (.77)	0.81 (.71)

Note. CBM-R = curriculum-based measurement of oral reading. Skew and kurtosis were within reasonable limits ($z \leq 2.00$) for all conditions. *N* values for second, third, fourth, fifth, and sixth grades were, respectively, 1790, 1885, 1913, 1942, and 1144 for general education; 111, 134, 170, 192, and 116 for special education.

from zero for all grade levels, meaning that both the slopes and intercepts at the district levels were identical. In addition, such a result supports evidence that the Level 2 LMM analysis introduced in the method section is appropriate for this study.

Results

Summaries of the means and standard deviations for CBM-R at each time point are presented in Table 1. As shown in Table 1, the average growth as measured with CBM-R was positive for all conditions across seasons (fall–winter, winter–spring), grade levels (second to sixth) and populations (general education, special education).

Comparison Between Linear Model and Piecewise Model

Results for the comparisons of linear and piecewise models are presented in Table 2. For students in general education, the piecewise model provided the best fit for tri-annual assessment data, as evidenced by the magnitudes of LRT, AIC, and BIC observed for the linear and piecewise models respectively. The χ^2 difference test between the two models also revealed that the piecewise model fit significantly better than the linear model.

For students in special education, on the basis of the LRT criterion, the piecewise model for all grade levels was viewed as the best-fitting model. Significant differences in LRTs between the two models, however, were found in the second through fourth grades. There were inconsistencies with respect to AIC and BIC criteria, which might indicate the linear model fit better in some cases within the sample of special education students. Although the results of AIC and BIC are noted, the LRT was the primary test/criterion used here.

Table 2
Deviation Between Linear and Piecewise Models for General Education Students and Special Education Students

	Linear Model			Piecewise Model			Differences (Linear-Piecewise Model)		
Grade	LRT	AIC	BIC	LRT	AIC	BIC	LRT	AIC	BIC
General education									
Second	46231.7	46239.7	46261.7	45542.6	45556.6	45595.0	689.1*	683.1	666.7
Third	49566.9	49574.9	49597.0	49357.7	49371.7	49410.5	209.2*	203.2	186.5
Fourth	49768.1	49776.1	49798.3	49520.1	49534.1	49573.0	248.0*	242.0	225.3
Fifth	50974.9	50982.9	51005.2	50866.6	50880.6	50919.6	108.3*	102.3	85.6
Sixth	29892.9	29900.9	29921.1	29873.6	29887.6	29922.9	19.3*	13.3	−1.8
Special education									
Second	2816.3	2824.3	2835.1	2802.6	2816.6	2835.6	13.7*	7.7	−0.5
Third	3501.8	3507.8	3516.5	3487.9	3495.9	3507.5	13.9*	11.9	9.0
Fourth	4394.9	4402.9	4415.4	4378.9	4392.9	4414.9	16.0*	10.0	0.5
Fifth	5090.6	5098.6	5111.6	5083.3	5097.3	5120.1	7.3	1.3	−8.5
Sixth	3048.2	3056.2	3067.2	3046.3	3060.3	3079.5	1.9	−4.1	−12.3

Note. LRT = likelihood ratio test; AIC = Akaike information criterion; BIC = Bayesian information criterion.
*$p < .001$.

Comparison Between the First Slope and Second Slope for General and Special Education

Table 3 presents estimates of both the first (fall–winter) and second (winter–spring) linear slope. In the cases of both students in general education and special education, all of the seasonal linear slopes were positive with magnitudes significantly different than zero ($p < .001$). That observation was consistent with the conclusion that reading achievement accelerated continuously throughout the academic year for all student samples. Although the mean rate of growth across seasons was consistently greater for students in general education as compared with students in special education, it was the first slopes that were most distinct. Further, the differences in the second slopes varied by grade level.

Additional analysis was conducted to examine whether the first linear slopes (fall) were significantly different from the second slopes (spring). Table 3 presents the differences between the two linear slopes along with the results of significance tests. Figure 1 depicts the slopes and model differences.

In the sample of general education students, statistically significant differences ($p < .001$) were observed between the first and second slopes in all grade-level conditions (Table 3). That is, the rate of growth for all grade levels dropped in the second part of the academic year. It is relevant to note that the magnitude of difference between the two slopes (seasonal effects) and standardized effect sizes consistently decreased as grade level increased. This might indicate that the seasonal effect in reading is larger in the lower primary grades than the upper primary grades. The power estimates for detecting seasonal effects for students in general education were very large for all grade levels (range = 0.96–1.00). This indicates a good chance of detecting a false null hypothesis (linear model) when the alternative model (piecewise model) is correct (Duncan et al., 2006).

In the sample of special education students, the average of the first slope (fall) was

Table 3
Comparison Between the First Slope and the Second Slope for General Education Students and Special Education Students

Grade	Intercept $\hat{\beta}_0$ (*SE*)	Fall/First Slope (Fall to Winter) $\hat{\beta}_1$ (*SE*)	Spring/ Second Slope (Winter to Spring) $\hat{\beta}_2$ (*SE*)	Difference Between Fall Slope and Spring Slopes $\hat{\beta}_2{}^a$ (*SE*)	Standardized Effect Sizes[b]	Power
General education						
Second	57.42* (.76)	1.71* (.02)	1.02* (.01)	0.69* (.02)	34.50	1.00
Third	82.31* (.85)	1.38* (.02)	0.97* (.02)	0.41* (.03)	13.60	1.00
Fourth	105.78* (.93)	1.19* (.02)	0.78* (.02)	0.41* (.12)	3.41	1.00
Fifth	124.43* (.89)	1.12* (.02)	0.83* (.02)	0.29* (.12)	2.41	1.00
Sixth	142.49* (1.21)	0.88* (.02)	0.74* (.02)	0.14* (.16)	0.87	0.96
Special education						
Second	34.02* (2.67)	1.17* (.08)	1.08* (.06)	0.09 (.10)	0.90	0.86
Third	50.85* (2.81)	1.09* (.07)	0.89* (.06)	0.20 (.10)	2.00	0.87
Fourth	75.50* (2.99)	0.99* (.05)	0.67* (.05)	0.32* (.08)	4.00	0.92
Fifth	86.28* (3.06)	0.90* (.05)	0.67* (.06)	0.23 (.09)	2.50	0.54
Sixth	106.98* (4.40)	0.69* (.07)	0.73* (.06)	−0.04 (.10)	0.40	0.11

Note. SE = standard error.
[a]The estimated coefficients of the second slope from an added growth model.
[b]Standardized effect sizes = $\hat{\beta}_2$/SE, which is a strategy developed by Satorra and Saris (1985), was used to calculate power.
*$p < .001$.

greater than the second slope (spring) in all grades except for sixth; however, those differences were not statistically significant except in the fourth-grade condition (Table 3). Moreover, the pattern of reduced seasonal effects for progressively higher grades did not generalize to the sample of special education students. Instead, the magnitude was greatest in fourth grade (difference in slopes = 0.32, $p < .001$), somewhat smaller in third (0.20, $p > .001$) and fifth (0.23, $p > .001$), and hardly perceptible in second (0.09, $p > .001$) and sixth grades (−0.04, $p < .001$). The range of standardized effect sizes and power was variable across grades. Based on the results, it appears that, in general, the observed slopes in the sample of special education students were relatively consistent across seasons—in all conditions but fourth grade.

Discussion

Many of the published estimates of CBM-R annual growth rates rely on the assumption that the pattern is consistent and linear throughout the academic year (Deno et al., 2001; Fuchs et al., 1993). Although there is evidence for seasonal effects within some studies (Ardoin & Christ, 2009; Fuchs et al., 1993; Graney et al., 2009), the phenomenon of seasonal effects are generally ignored within the professional literature. This study yielded (a) estimates for the magnitude CBM-R annual growth along with those within the fall and spring seasons, and estimates were derived for both general education and special education populations; (b) both linear and piecewise growth models; and (c) hypothesis tests to examine whether fall growth is typically more substantial than spring growth.

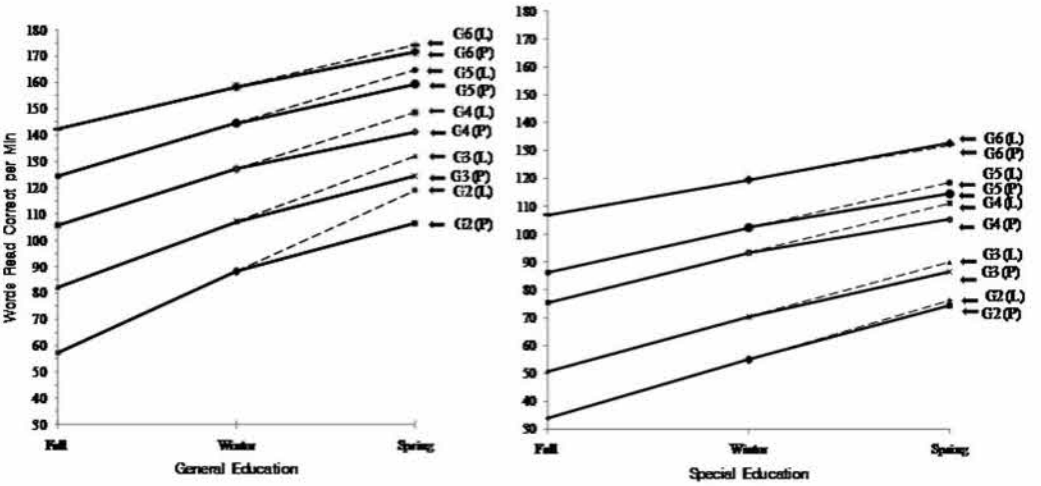


Figure 1. Differences for linear (L, solid line) and piecewise (P, dashed line) models of growth (G) for students in second through sixth grade within either the general (left) or special education population (right)

CBM-R growth was modeled and analyzed for five primary grades (second through sixth) and two student populations within and across seasons.

The results of this study support four general conclusions. Descriptive data make clear that, first, annual growth is greater among students in the general education population than for those in the special education population, especially in the fall season; and, second, more growth occurs in the early grades with less in the upper grades. Modeling and inferential analysis provide support for the third and fourth conclusions that follow. That is (third conclusion), there is a significant *seasonal effect* such that more growth occurs in the fall season than in the spring season for general education, but the magnitude of that effect declines with each progressive grade level. Finally, the seasonal effect is less pronounced among students in the special education population. In general, linear growth should not be assumed, seasonal effects should be expected—at least within the general education population—and piecewise models tend to fit better than linear models of growth.

Grade and Population Effects

Larger magnitudes of growth were observed for students in general education as compared to special education. This outcome was observed across five districts where CBM-R was used to screen, monitor, and evaluate instructional effects as part of a problem solving model. The systematic and ongoing use of CBM-R for these purposes did not establish equivalent growth rates between the student populations. Despite inconsistent growth rates across student groups, the rate of growth among students in the special education population was greater than observed in other studies (Deno et al., 2001; Fuchs et al., 1993). For example, the most recent estimates for average growth among students in the special education population were ~0.60 WRCM/week (Deno et al., 2001). That rate of growth was relatively consistent across grades. In contrast, the average growth estimates among students in the sample of students in special education used in this study were within the range of 0.67–1.17 WRCM/week (Table 3). Although cross-study comparisons cannot be used to establish that differences in systems-level practices (e.g., problem solving) and instructional practices were causally related to

the observed differences in growth rates, future research should investigate the phenomena more closely.

The results of this study converge with previous research (Deno et al., 2001; Fuchs et al., 1993) to support the conclusion that there is more CBM-R growth within the early primary grades. This phenomenon was consistent for both fall and spring seasons in both the general education and special education populations (Table 1). There was a steady decline for both general education and special education in weekly growth as grade level was increased. There was a corresponding decline in the magnitude of seasonal effects as grade levels increased. Results also support the conclusion that seasonal effects are more robust for students served within general education and less robust for students served within special education.

Seasonal Effect

The primary purpose of this research was to determine whether the rate of CBM-R growth is consistent across the year. Notwithstanding limitations, a seasonal effect was sufficiently robust and consistent across grades and student populations to warrant serious consideration. Estimates of weekly CBM-R growth within season are reported in Table 1. They are also depicted in Figure 1. Both visual analysis and statistical analysis of growth models support the conclusion that there was a *difference in the average rate of growth across seasons*, and that difference occurred in all grades, but was generally more robust in the lower grades of general education (Figure 1). With few exceptions, more growth occurred in the fall than in the spring, which is depicted as the differences in the slopes within Table 3.

The *magnitudes of seasonal effects were inconsistent* across grades and student populations. In general education, the difference between fall and spring slopes ranged from 0.14 to 0.69 per week, and all the differences were statistically significant with substantial effect sizes (Table 3). That effect is illustrated in Figure 1, which reveals that the interactions between seasonal effect and student population were most distinct in second grade. That is, the seasonal effect was most robust for the sample of students in second-grade general education classes, with a difference in growth from fall to winter of 0.69, and minor in comparison for students in second-grade special education classes, with a difference of 0.09.

The results of this study support the conclusion that growth differs between fall and spring seasons, which is inconsistent with a number of currently held views in the field. The consistency of the results across grades for general education begs the question of why this effect may be occur and what are the potential implications. There are a number of alternate hypotheses that offer some insight on the phenomenon and directions for future research.

Explanatory Hypotheses

This study was designed to examine *whether* seasonal effects are likely to occur in cases of CBM-R tri-annual assessment and challenge the assumption of linear growth. Indeed, the results do provide support for seasonal effects; however, this study was not designed to examine the cause of seasonal effects. That is left to future research to isolate the cause(s) and potential remedies to accelerate reading achievement within seasons with deficit growth. The result of both this study and Ardoin and Christ (2008) converge to provide evidence that the seasonal effect is likely to coincide with deficit levels of growth in the latter part of the academic year, or the spring season, especially for the students in the early primary general education population. It seems that changes in instructional conditions, or student motivation, might explain deficit rates of growth within the spring season unless there is an active intervention to prevent those changes. It seems that data-based decision making could function to reduce, or reverse, the pattern of seasonal effects that were replicated across this study and that of Ardoin and Christ.

Teacher expectations. There are a number of ways in which teacher expectations can have an influence on student behavior,

which might be conveyed through discrepant use of praise and punishment statements, content-relevant feedback, individualized attention, differentiated instruction, and variability in the opportunities to respond and engage in instructionally relevant practice (Good, 1981). Prior research yielded conclusions that there are indeed sometimes differences in teacher treatment of high- and low-achieving students over the course of the year, which affected the amount and type of instruction and feedback that students received at different points in the academic year (Good et al., 1980). Researchers concluded that teachers are more invested in influencing and shaping student behavior at the beginning of the school year than later in the year. Furthermore, the differentiation in teacher behavior may have sustained effects and become more influential as the school year progresses, resulting in lower achievement and growth later in the school year in low-achieving students. Other research also demonstrated the potential influence of teacher expectations, including expectation biases across ethnic groups (McKown & Weinstein, 2007) and ability groups (Rosenthal & Jacobson, 1966).

Classroom management. There is evidence that classroom management and organizational strategies are implemented with greater fidelity and effort in the early part of the academic year (Cameron, Connor, & Morrison, 2004). The same research provides evidence that that those educators who approach classroom management and organization with the greatest fervor early in the year are those same educators whose practices drop off most substantially later in the academic year. Improved levels of engagement and instructional gain are clearly associated with good classroom management and organization, and are among better predictors of student achievement (Brophy, 1987). Taken together, research provides a foundation to postulate that differential levels of classroom management and organization across seasons might influence student achievement. This effect is hypothesized to be the greatest at lower grade levels, as younger students require greater direction in organization practices (Cameron et al., 2004). The pattern of seasonal effects observed in this study coincides with a classroom management and organizational hypothesis.

Intervention to accelerate seasonal growth. The results of this study indicate that the seasonal effect with deficit growth in the spring season was most pronounced in the middle primary grades (second and third) and were minimized for the special education population. The results of Graney et al. (2009) indicate that the seasonal effect can function so that there is deficit growth in the fall season and accelerated growth in the spring. As noted in the Introduction, the data set that was analyzed by Graney et al. (2009) derived from a university and school partnership to implement a data-based decision making and evidence-based interventions. Those factors might have influenced the patterns of seasonal effects. There is a substantial literature base to support the conclusion that data-based decision making can function to accelerate academic growth (Carnine & Granzin, 2001; Fuchs & Fuchs, 1986; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Fuchs, Fuchs, Hamlett, & Whinnery, 1991; Graney et al., 2009), and there is substantial theoretical support for the use of evidence-based practices, which were both components of service delivery that were targeted by the partnership program. It might be that progress monitoring and data use accelerated learning and ameliorated the deficit growth associated with seasonal effects. Indeed, students who are served within special education have explicit goals and their progress is monitored. Although it is not possible to draw any final conclusions, it may well be that the lack of convergence in seasonal patterns between Graney et al. and other studies—along with minimal seasonal effects observed for the special education population—provides further evidence for the use of data-based decision making to improve growth during otherwise deficit periods.

Limitations and Future Research

The sample studied was entirely from five relatively ethnically homogeneous dis-

tricts in a single midwestern state. Future research should attempt to corroborate these findings across a more diverse sample of students and educational programs. Finally, the results of this study derive from data collected at three points in the academic year (fall, winter, spring). Future research should examine whether the same conclusions would result from the analysis of a more robust data set (e.g., weekly, monthly) and with more precise dates that define the data collection schedule. Graney et al. (2009) correctly criticize the methods of this study and that of Ardoin and Christ (2008) because the precise dates of data collection were not used in the analysis; rather, both studies rely on data that are centered by month and not day. Although the authors of this and the prior study believe the conclusions would be substantially similar with different centering methods or more precise dates, the criticism ought to be accounted for in future research. Future research should investigate the nature of growth across summative and formative measurement conditions. Research should also examine the effect on educational decisions when the potential for seasonal effects is ignored and consistent annual growth is assumed.

Implications for School-Based Decision Making

CBM-R was developed to guide routine instructional decisions and enhance instructional effects (Deno, 1986), and it has emerged as one of the primary methods of assessment to guide problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention models of service delivery (National Association of State Directors of Education, 2005). As described in the Introduction of this study, CBM-R is used to establish expectations for both the level and rate of reading achievement throughout the primary grades; therefore, the results of this study provide at least two clear implications. First, there is good reason to question the validity of linear growth estimates as criteria to guide progress monitoring (i.e., discrepant in rate) and dual-discrepancy decisions (i.e., discrepant in both level and rate; Fuchs, Fuchs, McMaster & Al Otaiba, 2003). Second, seasonal effects can be either admired as a phenomenon or targeted for intervention. These implications are addressed in what follows.

First, if CBM-R estimates of growth are used to evaluate instructional effects and response to intervention, then those who interpret data should be aware that the magnitude of growth might be less in the spring than in the fall. There are two weighty implications that derive from this study. First, a dual-discrepancy approach to special education eligibility, which compares the observed level and rate of achievement to criterion standards (Fuchs, Fuchs, McMaster & Al Otaiba, 2003), might function to overidentify students during deficit seasons of growth and underidentify students during accelerated seasons of growth. If the data follow the pattern observed in this study, then it is likely that the dual-discrepancy model might underidentify students in the fall season and overidentify students in the spring season. That consequence is counter to the values of early intervention and prevention inherent within problem solving and response to intervention. Researchers and practitioners should pay close attention to the patterns of growth and the potential for seasonal effects to influence decisions at the local level. Accelerated rates of growth should be expected in the fall.

Second, the pattern of seasonal effects that are observed across populations and studies provide preliminary evidence that goal setting, progress monitoring and data-based decision making can minimize, eliminate or reverse the pattern of seasonal effects (e.g., Graney et al., 2009). Although future research is necessary, it is a working thesis that the systematic use of data to evaluate instructional effects and select effective instructional strategies effectively accelerates learning and prevents problems at the individual, group, and systems level. After all, data-based decision making is a central tenet of problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention (National Association of State Directors of Education, 2005); however, data must be both systematically collected *and* sys-

tematically used to establish effects (Fuchs et al., 1992, 1991; Fuchs & Fuchs, 1986). The mere act of data collection without data use typically fails to establish effects (Fuchs & Fuchs, 1986). The explanatory hypotheses proposed in the previous section establish that deficit growth in the spring season might be remediated at the systems level if data are systematically collected and systematically used. If detected, this problem (seasonal effect) is ripe for problem analysis at the local level. Deficit growth within a single season should not be admired as an acceptable phenomenon, but, instead, it should be targeted for analysis and intervention.

Christ (2008) defined problem analysis as the "collection, summary, and use of information to systematically test, reject, or verify relevant hypotheses to establish problem solutions" (p. 159). That is, deficit rates of growth in the spring season might be remediated if expectations remain high and data are used to evaluate and maintain high standards for instructional effects. There is no reason to assume that the seasonal effect is inherent to the educational system. Instead, hypotheses related to the causal and maintaining variables should be developed—such as those presented earlier—and tested through intervention to remediate the problem. It is likely that the pattern of seasonal effects is causally related to what occurs in the classroom; and what occurs in the classroom can be influenced by the systematic used of data to evaluate of response to intervention for universal, supplemental, and intensive tiers of service delivery. This study provides evidence of seasonal effects as a potential target for intervention at the system, grade, or classroom level.

References

AIMSweb. (2002). *Reading curriculum-based measurement (R-CBM) standard oral reading fluency assessment passages.* Retrieved October 2003, from http://www.aimsweb.com

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum based measurement slope estimate using data from tri-annual universal screenings *School Psychology Review, 37,* 109–125.

Bollman, K. A., Silberglitt, B., & Gibbons, K. A. (2007). The SCRED Model: Incorporating systems-level organization and a multi-tiered problem-solving process for intervention delivery. In M. Burns, S. Jimerson, & A. VanDerHayden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 321–332). New York: Springer.

Brophy, J. (1987). Synthesis of research on strategies for motivating students to learn. *Educational Leadership, 45*(2), 40–48.

Cameron, C. E., Connor, C. M., & Morrison, F. J. (2004). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology, 43,* 61–85.

Carnine, D., & Granzin, A. (2001). Setting learning expectations for students with disabilities. *School Psychology Review, 30,* 466–472.

Christ, T. J. (2008). Best practices in problem analysis. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 159–176). Bethesda, MD: National Association of School Psychologists.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74,* 68.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219–232.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15,* 358–374.

Deno, S. L. (2002). Problem solving as best practices. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 37–56). Bethesda, MD: National Association of School Psychologists.

Deno, S. L. (2005). Problem solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 10–42). New York: Guilford Press.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30,* 507–524.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). An introduction to latent variable growth curve modeling: Concepts, issues, and applications (2nd ed.). Mahwah, NY: Lawrence Erlbaum Associates.

Flora, D. B. (2008). Specifying piecewise latent trajectory models for longitudinal data. *Structural Equation Modeling, 15,* 513–533.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53,* 3–14.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10,* 43–52.

Fuchs, L. S., Fuchs, D., & Stecker, P. M. (1989). Effects of curriculum-based measurement on teachers' instructional planning. *Journal of Learning Disabilities, 22,* 51–59.

Fuchs, L. S., & Shinn, M. R. (1989). Writing CBM IEP objectives. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 131–152). New York: Guildford Press.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58,* 436–450.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22,* 27–48.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Whinnery, K. (1991). Effects of goal line feedback on level, slope, and stability of performance within curriculum-based measurement. *Learning Disabilities Research and Practice, 6,* 66–74.

Fuchs, D., Fuchs, L. S., McMaster, K. N., & Al Otaiba, S. (2003). Identifying children at risk for reading failure: Curriculum-based measurement and the dual-discrepancy approach. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 431–449). New York: Guilford Press.

Good, T. L. (1981). Teacher expectations and student perceptions: A decade of research. *Educational Leadership, 38*(5), 415–422.

Good, T., Cooper, H., & Blakey, X. (1980). Classroom interaction as a function of teacher expectations, student sex and time of year. *Journal of Educational Psychology, 72*, 378–385.

Graney, S. B., Missall, K. N., Martinez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School* Psychology, 47, 121–142.

Howe, K. B., & Shinn, M. M. (2002). Standard *reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Retrieved October 2004, from http://www.aimsweb.com

Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education for Students Placed at Risk, 9,* 169–188.

McKown, C., Weinstein, R. S. (2007). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*, 235–261.

National Association of State Directors of Education. (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: Author.

National Reading Panel. (2000). *Report of the National Reading Panel Subgroups: Teaching children to read* (No. NIH 00–4754). Bethesda, MD: National Institute of Child Health and Human Development.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports, 19,* 115–118.

Satorra, A., & Saris, W. E. (1985). The power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50*, 83–90.

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education, 34,* 164–172.

Shinn, M. R. (1988). Development of curriculum-based local norms for use in special education decision-making. *School Psychology Review, 17,* 61–80.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children.* New York: Guildford Press.

Shinn, M. (2008). Best practices for using curriculum based measurement in a problem solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 243–262). Bethesda, MD: National Association of School Psychologists.

Stage, S. A. (2001). Program evaluation using hierarchical linear modeling with curriculum-based measurement reading probes. *School Psychology Quarterly, 16,* 91–112.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41,* 85–120.

Date Received: May 30, 2007
Date Accepted: April 21, 2010
Action Editor: Tanya Eckert ■

Theodore J. Christ, PhD, is an associate professor in the Educational Psychology Department at the University of Minnesota in Minneapolis—Saint Paul. He is the codirector of the Research Institute for Problem Solving. He is engaged in research and training to evaluate, develop, and enhance the tools and procedures that enable data-based decision making and problem solving within school settings. Interests include progress monitoring, curriculum-based measurement, direct behavior ratings, and computer adaptive testing.

Benjamin Silberglitt completed his PhD in Educational Psychology at the University of Minnesota at Minneapolis—Saint Paul. His areas of interest include student assessment school organization and system-level change, problem solving, response to intervention, and data-based decision making. He is director of Software Applications for TIES, a technology cooperative of 39 Minnesota school districts.

Seungsoo Yeo, PhD, is a researcher at Korean Educational Development Institute. He received his PhD from the University of Minnesota at Twin Cities in 2008. His primary research interests include curriculum-based measurement, response to intervention, and learning disabilities. His methodological work has focused on latent growth modeling and multilevel meta-analysis.

Damien C. Cormier, MA, is a school psychology doctoral student in the Educational Psychology Department at the University of Minnesota. He is a graduate research assistant at the National Center on Educational Outcomes. His research interests are in the areas of large-scale assessment policy and practice, measurement issues in individual assessment, high-functioning autism, and cultural and linguistic biases in assessment.

Copyright of School Psychology Review is the property of National Association of School Psychologists and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

I was mistaken when I indicated that the Reading Fluency Monitor was developed by Read Naturally. It was actually developed by the RMC Research Corporation out of Portsmouth, New Hampshire.

Sorry for the misunderstanding,

- Dave

David Heistad, Ph.D., Executive Director
Research and Evaluation Department
Bloomington Public Schools - ISD 271
1350 W. 106th Street
Bloomington, MN 55431
ph: 952-681-6485
cell: 952-250-5971

Karla Ramy

From:Karla Ramy
Sent:24 Oct 2013 12:42:19 +0000
To:What Works
Subject:Read Naturally - Tom Ihnot
Attachments:Christ Silberglitt Yeo Cormier (2010).pdf, Heistad.docx
Dear Neil Seftor,

In an August 26, 2013 letter from What Works Clearinghouse (WWC), Jill Constantine stated:

> *Order of findings.* We appreciate and agree with your suggestion of July 17, in which you propose that the WWC list findings in order of their strength, consistent with the presentation on the web. Therefore, we will revise the order of findings listed in the text on the front page of the report, the associated table, and the website summary for the three reports with findings in more than one domain. We anticipate that the revised intervention reports will be posted to the web by the end of September. (Reference :QR2010013, WWC2304)

I look forward to this change being executed as soon as possible. Please email me a verification that the change in order of findings will read:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

Also, I received an October 18, 2013 email (Reference: QR2013002) responding to my concerns about the Heistad and Arvans studies. Thank you for your response. Concerning the Heistad study, you stated that the measure was over-aligned with the intervention as words used in the Read Naturally curriculum also appear in the Reading Fluency Monitor. I do not understand this reasoning. Please explain it to me in more detail. I am concerned that the following facts are not being properly evaluated:

- Read Naturally passages were developed by Read Naturally staff as curriculum.
- The Reading Fluency Monitor was developed by RMC Research Corporation, located in Portsmouth, New Hampshire, as a benchmark assessment tool.
- Read Naturally curriculum passages and the Reading Fluency Monitor are completely different tools developed by different companies.
- Dr. Heistad clearly states that he made a mistake in the test instruments section of his study. His verification of this is attached.

Including Dr. Heistad's statistically significant gains in oral reading fluency for third-grade students across multiple schools would change the effectiveness rating in the Beginning Reading section for Read Naturally from mixed to positive effects.

The Arvans study clearly shows that Read Naturally programs are research proven when it comes to oral reading fluency gains. The group using the Read Naturally strategy had a large effect size of .81. When you conduct a short eight-week study, it is an accepted practice to do an extrapolation for a school

year. (Christ – Attached) What Works Clearinghouse should consider changing its policy on extrapolations.

I look forward to a prompt listing and announcement of the new wording on Read Naturally's effectiveness rating for the Beginning Reading Section. Also, I would appreciate a reevaluation of the over-alignment issue with the Heistad study and further consideration of the Arvans study.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

School Psychology Review,
2010, Volume 39, No. 3, pp. 447–462

Curriculum-Based Measurement of Oral Reading: An Evaluation of Growth Rates and Seasonal Effects Among Students Served in General and Special Education

Theodore J. Christ
University of Minnesota

Benjamin Silberglitt
Technology and Information Educational Services

Seungsoo Yeo
Korean Educational Development Institute

Damien Cormier
University of Minnesota

Abstract. Curriculum-based measurement of oral reading (CBM-R) is often used to benchmark growth in the fall, winter, and spring. CBM-R is also used to set goals and monitor student progress between benchmarking occasions. The results of previous research establish an expectation that weekly growth on CBM-R tasks is consistently linear throughout the academic year. The patterns of CBM-R growth were examined for a large sample of students (N = 3808) from both general education and special education populations in second to sixth grades. Results support four general conclusions: (a) annual growth is more substantial within the general education population; (b) growth is more substantial in earlier elementary grades; (c) more growth occurs in the fall than the spring season (i.e., seasonal effect), especially within the early primary general education population; and (d) the seasonal effect is less pronounced within the special education population. Estimates of growth within and across seasons are presented and implications are discussed.

Curriculum-based measurement (CBM) is used to index annual student growth across the primary grades. Procedures and measurement metrics are developed for mathematics, spelling, written expression, and reading. CBM oral reading (CBM-R) rate is the most researched and well established of those available procedures (Wayman, Wallace, Wiley, Ticha, & Espin, 2007). CBM-R is used as an index of reading rate and fluency, which are identified as critical skills to target for instruction and intervention within the early stages of

Correspondence regarding this article should be addressed to Theodore J. Christ, 56 East River Road, Minneapolis, MN 55455; E-mail: tchrist@umn.edu

Copyright 2010 by the National Association of School Psychologists, ISSN 0279-6015

reading development (National Reading Panel, 2000; National Research Council, 1998). CBM-R outcome also functions as a robust indicator of overall reading development throughout the primary grades (Wayman et al., 2007).

CBM-R progress monitoring data are collected and used to evaluate instructional effects, determine when to modify the instruction, and evaluate response to intervention (Deno, 1985, 1986). Teachers who use progress monitoring data are likely to use more specific and measurable goals, rely more substantially on data to guide instruction, and modify instruction more frequently (Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, & Stecker, 1989). Progress monitoring applications of CBM-R have the potential to confer the most substantial benefit to students. Progress monitoring can be conceptualized as either interim benchmark/screening assessments when data are collected three to four times per year, or time series continuous assessments when data are collected daily or weekly.

Both interim and time series CBM-R data are plotted graphically and evaluated against goals and goal lines. Fuchs and Shinn (1989) recommended that the rate of expected growth should derive from local normative performance so that individual student achievement can be compared against the local normative performance of grade-level peers. A goal line is established by connecting the CBM-R data point for the observed level of performance from Week 1 (initial performance) to the expected/typical level of performance at Week 36 (ending performance). Once graphed, the goal line provides a graphic trajectory of expected growth for the academic year. Ongoing time series data are then plotted on the graph and evaluated against the goal line. For example, if a student's initial level of CBM-R performance in September was 20 words read correctly per minute (WRCM) and in May that student was expected to read 60 WRCM, then the expected annual growth is 40 WRCM across 8 months, or 32 weeks. That translates to ~1.25 WRCM per week of growth (1.25 WRCM = 40 WRCM/32 weeks). The goal line and trajectory of expected growth is based on the assumption of a monotonic and linear trend.

Monotonic Linear Growth

Standards of expected performance within and across the year provide the foundation to establish instructional goals and goal lines. There are two influential studies that established standards of weekly CBM-R growth (Deno, Fuchs, Marston, & Shin, 2001; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993), which have application in both research and practice. Fuchs et al. (1993) analyzed a CBM-R data set comprised of 103 students in general education and 14 students in special education. Each student was assessed weekly for an entire academic year using CBM-R procedures. Although some students evidenced a negatively accelerating annual growth, the researchers concluded that annual growth is, on average, monotonic and linear. They also observed a negatively accelerating trend across grades. That is, the rate of CBM-R growth was positive in all grades, but the rate of growth declined in each successive grade. The resulting analysis yielded grade-specific estimates of standard and ambitious rates of growth—and *assumed* a pattern of monotonic linear growth. Respectively, the growth, determined by the number of WRCM, were 1.5 and 2.0 in first and second grade, 1.0 and 1.5 in third, 0.85 and 1.1 in fourth, 0.5 and 0.8 in fifth, and 0.3 and 0.63 in sixth.

Deno et al. (2001) analyzed a large data set comprised of 2675 students in general education and 324 in special education classes. The student sample was from school districts across the country. The researchers relied on the *assumption* of monotonic linear annual growth. Ordinary least squares regression was used to calculate weekly growth. The study yielded useful estimates of observed and expected growth across grades and populations. The authors estimated that beginning readers in the general education population can be expected to improve 2 WRCM/week until they achieve 30 WRCM; thereafter, students in the general education population can be expect to improve at least 1 WRCM/week. Researchers observed that students in the special education

population demonstrated substantially less growth, which approximated 0.5–0.8 WRCM/week, unless they are provided with robust evidence-based instruction, which might improve the observed rate of growth to 1.39 WRCM/week.

Prior studies provide useful estimates of typical, ambitious, and expected rates of CBM-R growth; however, those estimates depend on the assumption that growth is monotonic and linear across the academic year. That assumption should be examined more closely to determine its veracity. The prior studies also rely on analyses that neglect the inherent nesting of students within classrooms and schools, which establish potential bias in the linear and quadratic functions derived.

Improved Modeling of Annual and Seasonal Growth

Research is necessary to evaluate alternate models of annual growth and test the predominant assumption that annual growth is monotonic and linear. This study will evaluate both linear (constant) and piece wise (nonconstant) models of growth from fall to winter to spring; moreover, this study will employ the superior approach of linear mixed model (LMM) rather than that of ordinary least squares used in prior research (Deno et al., 2001; Fuchs et al., 1993).

Ordinary least squares only estimates the fixed effects, which assumes that all cases have identical parameter estimates for intercept and growth. This is a limitation when modeling CBM-R growth because it is likely to be an erroneous assumption in most cases. The results of published research provide evidence for substantial magnitudes of variance in both intercept and growth among students in a sample (Shin, Deno, & Espin, 2000; Stage, 2001). It is likely that the intercept and growth rates should be treated as random effects in many instances. LMM can be used to estimate both the fixed effects and the random effects, which are the intercepts and growth rates for individual student cases in the sample. LMM takes into account the variances within and across individuals in the sample. Finally, ordinary least squares requires that missing data are removed casewise whereas LMM can handle missing data, which is likely to yield less biased estimates of intercept and slopes while enhancing power (Duncan, Duncan, & Strycker, 2006). This approach facilitates analysis of both seasonal effects, as pursued by two previous studies (Ardoin & Christ, 2008; Graney, Missall, Martinez, & Bergstrom, 2009), and compares those effects to that of linear growth as assumed in other studies (Deno et al., 2001; Fuchs et al., 1993).

The calendar year was divided into three seasons for the purpose of this study. The seasons were defined by those periods between standard tri-annual assessment occasions: fall, winter and spring. The duration between fall and winter assessments defines the fall season, that between winter and spring defines the spring season, and that between spring and fall defines the summer season. If growth is consistent and linear throughout the calendar year, then there is no *seasonal effect*. Conversely, if growth is not consistent and linear throughout the calendar year, then there is a seasonal effect whereby there is more growth observed for some seasons than for others.

The summer season is the most distinct for the majority of students because they are not exposed to formal instruction. The rate of achievement is likely to decline in the summer months (Kim, 2004) unless formal instruction is provided through summer school programs (Stage, 2001). A seasonal effect in the summer is expected because the conditions for learning are substantially different from that which occurs in the spring and fall seasons. The effect, however, does not seem specific to those differences between summer and the academic year. Ardoin and Christ (2008) observed statistically significant differences in CBM-R growth between the fall and spring seasons. That is, among a sample of second-grade students there was more growth in the fall than in the spring season. That result replicated across three combinations of CBM-R passages sampled from the Dynamic Indicators of Basic Early Literacy Skills. A subsequent study aimed to replicate and extend those findings

with a sample of third- through fifth-grade students across a variety of CBM curriculum domains (Graney et al., 2009). However, Graney et al. (2009) concluded that there was more CBM-R growth in the spring season than in the fall season, failing to replicate the work of Ardoin and Christ. Graney et al. speculated that the failure to replicate findings across studies might relate to the specification of data collection schedules, sample characteristics, issues of instrumentation, or the distribution of instructional intensity across the academic year. Although it is difficult to make a single attribution, it should be noted that the student sample used by Graney et al. derived from a school–university partnership program that included features to promote data use and evidence-based interventions. Indeed, that program might influence the pattern of student growth; however, additional research and replications are necessary to help determine whether specific patterns for seasonal effects should be expected.

There are limitations to the analytic procedures used in the two prior studies. Both generated estimates of growth from difference scores across seasons. That is, growth was estimated as the difference between performance at Time 1 and Time 2, which was then divided by the number of intervening weeks to yield an estimate of weekly growth. Although this method is intuitive, and it was useful for the preliminary work to examine seasonal effects, the method of difference scores to estimate change has long been criticized as an unreliable method of analysis (Cronbach & Furby, 1970) and may have contributed to the inconsistent findings. The analytic procedures used in this study improve on those of prior studies that examined seasonal effects and modeled annual growth. The analytic procedures used in this study will compare the hypothesized seasonal effect directly with that of linear growth.

Purpose

The purposes of this study were to (a) estimate the rate of growth within and across seasons for general education and special education populations; (b) evaluate whether the rate of CBM-R growth is likely to be consistent (i.e., linear model) or inconsistent (i.e., piecewise model) throughout the academic year; and (c) test the hypothesis that the seasonal effect is typically characterized by more robust growth in the fall than for the spring season. That hypothesis is counter to the assumptions inherent to the linear models as applied in many previous studies and it is counter to the findings of Graney et al. (2009).

Method

Participants and Setting

A total of 4824 students (52% male, 48% female) in the second through sixth grades were included in the sample. The ethnicity breakdown of the sample was 94% White, not Hispanic; 2% Native American; 2% Asian or Pacific Islander; 1% Hispanic; and 1% Black, not Hispanic. Participants were enrolled in seven elementary schools, located within five school districts in the rural and suburban Midwest. The sample consisted of 31% students receiving free or reduced-price lunch, and 8% students receiving special education services.

As part of an ongoing benchmark assessment plan and response to intervention implementation (see Bollman, Silberglitt, & Gibbons, 2007, for additional information), all students in participating schools were assessed using CBM-R every fall, winter, and spring of every year across the grade levels studied. All students for whom relevant data were collected participated in the study. Data were gathered between the years of 2001 and 2005 for each student in the sample. Data from across multiple years were available for many of the students in the sample; thus many students participated in the study at multiple grade levels.

Measures

The CBM-R measures were developed or selected to be curriculum-independent, grade-level appropriate, and of equivalent difficulty across probes within each grade level. Each probe contained approximately 250

words typed on an 8 by 10 piece of paper in a grade-appropriate font (range = 12–14 point with larger print for the second and third grades). Probes were selected from a standardized set of commercially available published passages (AIMSweb; Howe & Shinn, 2002). The technical manual for AIMSweb passages (Howe & Shinn, 2002) and peer-reviewed research (Wayman et al., 2007) provide evidence that test–retest reliability and alternate-form reliabilities approximate or exceed .90. Criterion-related validity coefficients typically approximate the .70–.90 range. The same set of CBM-R probes were used across all districts for all 4 years of the study.

Procedure

CBM-R data were collected in 4-month intervals in the fall, winter, and spring of each academic year. Standardized administration and scoring procedures were implemented for each administration, which was conducted by trained school-based personnel and high school honors students. These individuals were trained to administer three successive CBM-R probes and score WRCM using standardized procedures that have appeared in the published literature (Howe & Shinn, 2002; Shinn, 1988, 1989). Administration occurred outside of the classroom to minimize distractions, in a separate room or quiet hallway while the administrator and student were seated at desks or in chairs.

Training included a brief instructional session that lasted approximately 1 hr. Each training session was followed by an assessment of competency that included practice and evaluation of scoring procedures. No formal criterion for scoring was established until the final years of the study. In 2003–2004 and 2004–2005, each scorer was required to come within 2 WCPM of the correct score on three consecutive videotaped assessments. Although data were not maintained, few scorers required more than one opportunity to meet criterion. Finally, inter-rater reliability was not evaluated; however, CBM-R procedures have established high levels of inter-rater reliability that appear consistently within the published literature. A survey of 10 arbitrarily selected CBM-R studies that were published between 1995 and 2005 reported inter-rater reliabilities that were consistently above .95.

Analytical Plan

A LMM analysis was performed to investigate evidence of seasonal effects for CBM-R. The present LMM analysis considered two hypothesized growth models: linear and piecewise (Duncan et al., 2006; Flora, 2008). The linear model assumes a static growth rate across time points—meaning that regardless of change over time, the estimated growth rate from the linear model is identical to the sample's growth rate. In contrast, the piecewise model assumes that the developmental trajectory is not consistent across time, making it possible to provide two different linear trends. Thus, the piecewise model takes into account a possible nonlinear trajectory for CBM-R. Based on the three CBM-R data points used in this study, the piecewise model can be parameterized into two levels of analyses. First, Level 1 can be expressed as

$$y_{it} = \beta_{0i} + \beta_{1i}\lambda_{1t} + \beta_{2i}\lambda_{2t} + \varepsilon_{it} \quad (1)$$

where y_{it} is the outcome of CBM-R for subject i at time t, β_{0i} is the value of CBM-R for subject i in fall, β_{1i} is the first linear trend of subject i from fall to winter, λ_{1t} is the value of the linear trend at time t (Week 0 and Week 18), β_{2i} is the second linear trend of subject i from winter to spring, λ_{2t} is the value of the linear trend at time t (Week 18 and Week 36), and ε_{it} is unique measurement error for subject i at time t.

Level 2 can be described as

$$\beta_{0i} = \beta_0\zeta_{0i} \quad (2)$$

$$\beta_{1i} = \beta_1\zeta_{1i} \quad (3)$$

$$\beta_{2i} = \beta_2\zeta_{2i} \quad (4)$$

where β_0 represents the average initial status for all subjects, β_1 refers to the average of the first linear growth rate for all subjects, and β_2

refers to the average of the second slope trend for all subjects. ζ_{0i}, ζ_{1i}, and ζ_{2i} refer to the random deviation of the initial status, the first growth rate, and the second growth rate, respectively. As a result, the piecewise model presented earlier assumes that the second slope is statistically different from the first slope, indicating that there is a seasonal effect for CBM-R throughout the academic year.

The first step in this analysis was to compare the two models (linear and piecewise models) using fit indices, which included the likelihood ratio test (LRT), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The three fit indices are considered to be "goodness-of-fit" indices, where the smallest fit statistics are viewed as the best model. If results from these goodness-of-fit indices reveal that the linear model is better than the piecewise model, then the slope is the same from fall to spring and there is no evidence of a significant seasonal effect for CBM-R across the academic year. Conversely, if goodness-of-fit indices are better for the piecewise model than the linear model, then we conclude that the slopes are different from fall to winter and from winter to spring and there is evidence of a seasonal effect across the academic year. If the piecewise model is the better fit, then an additional model is needed to estimate the extent to which the first slope differs from the second slope.

The additional model, which is called *added growth model,* would be identical to the piecewise model described earlier, except for the supplementary time factor loading matrix (Flora, 2008). For example, in the piecewise model, the time factor loading can be described as

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 18 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (5)$$

where the first column is set at 1, meaning that the outcome of the fall CBM-R score (intercept) does not change across time. The second column refers to the first linear time metric indicating change in CBM-R score from fall to winter (Flora, 2008). Similarly, the third column represents change in CBM-R score from winter to fall.

For the added growth model, the time factor loading can be expressed as:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 36 & 18 \end{bmatrix} \begin{matrix} Week & 0 \\ Week & 18 \\ Week & 36 \end{matrix} \quad (6)$$

The matrix in Equation 6 is identical to the matrix in Equation 5, with the exception of the second column with 0, 18, and 36. The different time factor loadings for the additional model is applied to examine the extent to which linear change in CBM-R from fall to winter increases or decreases compared to that from winter to spring. In other words, the added growth model analysis makes it possible to contrast the first linear slope and the second linear slope. The results from the additional analysis can also be viewed as an effect size estimate between the first and second linear slope (Flora, 2008).

Data Preparation

Overall, the rates of missing data for fall, winter, and spring seasons were less than 4% for all grade levels. All LMM analyses were conducted using the SAS MIXED program with full maximum likelihood estimation that automatically conducts all adjustments for missing data. Monthly CBM data used in this study were nested within districts. Therefore, a three-level LMM was used to satisfy an assumption that data collected in this study might not be independent among school districts. An alpha level for all statistical tests in this study was set at .001.

Model assumptions were evaluated, which included a test of normality using Kolmogorov-Smirnove statistics, skew, and kurtosis, and homoscedasticity of residuals by plotting raw residuals at each level (Levels 1 and 2). No violations were detected. The Level 3 LMM analysis was used to test whether CBM data used in the study were nested within school districts. As a result, variations in the slopes and intercepts between school districts were not significantly different

Table 1
Descriptive Statistics for the Level and Change of CBM-R

	Seasonal Benchmark Levels			Weekly Change	
Population	Fall *M* (*SD*)	Winter *M* (*SD*)	Spring *M* (*SD*)	Fall Season (Fall–Winter) *M* (*SD*)	Spring Season (Winter–Spring) *M* (*SD*)
General education					
Second	57 (32)	88 (34)	107 (35)	1.71 (.73)	1.02 (.64)
Third	82 (37)	107 (38)	125 (40)	1.38 (.79)	0.97 (.75)
Fourth	106 (38)	128 (39)	141 (41)	1.19 (.72)	0.78 (.70)
Fifth	125 (40)	145 (41)	160 (42)	1.12 (.76)	0.83 (.75)
Sixth	142 (41)	158 (42)	171 (43)	0.89 (.71)	0.74 (.75)
Total	99 (47)	122 (45)	137 (46)	1.26 (.74)	0.87 (.72)
Special education					
Second	33 (28)	55 (35)	74 (38)	1.17 (.81)	1.08 (.63)
Third	52 (33)	72 (39)	88 (41)	1.09 (.81)	0.89 (.78)
Fourth	77 (40)	96 (43)	108 (45)	0.99 (.72)	0.67 (.65)
Fifth	89 (43)	106 (45)	118 (47)	0.90 (.76)	0.67 (.79)
Sixth	106 (47)	119 (50)	132 (51)	0.70 (.72)	0.73 (.69)
Total	72 (45)	89 (48)	104 (49)	0.97 (.77)	0.81 (.71)

Note. CBM-R = curriculum-based measurement of oral reading. Skew and kurtosis were within reasonable limits ($z \leq 2.00$) for all conditions. *N* values for second, third, fourth, fifth, and sixth grades were, respectively, 1790, 1885, 1913, 1942, and 1144 for general education; 111, 134, 170, 192, and 116 for special education.

from zero for all grade levels, meaning that both the slopes and intercepts at the district levels were identical. In addition, such a result supports evidence that the Level 2 LMM analysis introduced in the method section is appropriate for this study.

Results

Summaries of the means and standard deviations for CBM-R at each time point are presented in Table 1. As shown in Table 1, the average growth as measured with CBM-R was positive for all conditions across seasons (fall–winter, winter–spring), grade levels (second to sixth) and populations (general education, special education).

Comparison Between Linear Model and Piecewise Model

Results for the comparisons of linear and piecewise models are presented in Table 2. For students in general education, the piecewise model provided the best fit for tri-annual assessment data, as evidenced by the magnitudes of LRT, AIC, and BIC observed for the linear and piecewise models respectively. The χ^2 difference test between the two models also revealed that the piecewise model fit significantly better than the linear model.

For students in special education, on the basis of the LRT criterion, the piecewise model for all grade levels was viewed as the best-fitting model. Significant differences in LRTs between the two models, however, were found in the second through fourth grades. There were inconsistencies with respect to AIC and BIC criteria, which might indicate the linear model fit better in some cases within the sample of special education students. Although the results of AIC and BIC are noted, the LRT was the primary test/criterion used here.

Table 2
Deviation Between Linear and Piecewise Models for General Education Students and Special Education Students

	Linear Model			Piecewise Model			Differences (Linear-Piecewise Model)		
Grade	LRT	AIC	BIC	LRT	AIC	BIC	LRT	AIC	BIC
General education									
Second	46231.7	46239.7	46261.7	45542.6	45556.6	45595.0	689.1*	683.1	666.7
Third	49566.9	49574.9	49597.0	49357.7	49371.7	49410.5	209.2*	203.2	186.5
Fourth	49768.1	49776.1	49798.3	49520.1	49534.1	49573.0	248.0*	242.0	225.3
Fifth	50974.9	50982.9	51005.2	50866.6	50880.6	50919.6	108.3*	102.3	85.6
Sixth	29892.9	29900.9	29921.1	29873.6	29887.6	29922.9	19.3*	13.3	−1.8
Special education									
Second	2816.3	2824.3	2835.1	2802.6	2816.6	2835.6	13.7*	7.7	−0.5
Third	3501.8	3507.8	3516.5	3487.9	3495.9	3507.5	13.9*	11.9	9.0
Fourth	4394.9	4402.9	4415.4	4378.9	4392.9	4414.9	16.0*	10.0	0.5
Fifth	5090.6	5098.6	5111.6	5083.3	5097.3	5120.1	7.3	1.3	−8.5
Sixth	3048.2	3056.2	3067.2	3046.3	3060.3	3079.5	1.9	−4.1	−12.3

Note. LRT = likelihood ratio test; AIC = Akaike information criterion; BIC = Bayesian information criterion.
*$p < .001$.

Comparison Between the First Slope and Second Slope for General and Special Education

Table 3 presents estimates of both the first (fall–winter) and second (winter–spring) linear slope. In the cases of both students in general education and special education, all of the seasonal linear slopes were positive with magnitudes significantly different than zero ($p < .001$). That observation was consistent with the conclusion that reading achievement accelerated continuously throughout the academic year for all student samples. Although the mean rate of growth across seasons was consistently greater for students in general education as compared with students in special education, it was the first slopes that were most distinct. Further, the differences in the second slopes varied by grade level.

Additional analysis was conducted to examine whether the first linear slopes (fall) were significantly different from the second slopes (spring). Table 3 presents the differences between the two linear slopes along with the results of significance tests. Figure 1 depicts the slopes and model differences.

In the sample of general education students, statistically significant differences ($p < .001$) were observed between the first and second slopes in all grade-level conditions (Table 3). That is, the rate of growth for all grade levels dropped in the second part of the academic year. It is relevant to note that the magnitude of difference between the two slopes (seasonal effects) and standardized effect sizes consistently decreased as grade level increased. This might indicate that the seasonal effect in reading is larger in the lower primary grades than the upper primary grades. The power estimates for detecting seasonal effects for students in general education were very large for all grade levels (range = 0.96–1.00). This indicates a good chance of detecting a false null hypothesis (linear model) when the alternative model (piecewise model) is correct (Duncan et al., 2006).

In the sample of special education students, the average of the first slope (fall) was

Table 3
Comparison Between the First Slope and the Second Slope for General Education Students and Special Education Students

Grade	Intercept $\hat{\beta}_0$ (*SE*)	Fall/First Slope (Fall to Winter) $\hat{\beta}_1$ (*SE*)	Spring/ Second Slope (Winter to Spring) $\hat{\beta}_2$ (*SE*)	Difference Between Fall Slope and Spring Slopes $\hat{\beta}_2{}^a$ (*SE*)	Standardized Effect Sizes[b]	Power
General education						
Second	57.42* (.76)	1.71* (.02)	1.02* (.01)	0.69* (.02)	34.50	1.00
Third	82.31* (.85)	1.38* (.02)	0.97* (.02)	0.41* (.03)	13.60	1.00
Fourth	105.78* (.93)	1.19* (.02)	0.78* (.02)	0.41* (.12)	3.41	1.00
Fifth	124.43* (.89)	1.12* (.02)	0.83* (.02)	0.29* (.12)	2.41	1.00
Sixth	142.49* (1.21)	0.88* (.02)	0.74* (.02)	0.14* (.16)	0.87	0.96
Special education						
Second	34.02* (2.67)	1.17* (.08)	1.08* (.06)	0.09 (.10)	0.90	0.86
Third	50.85* (2.81)	1.09* (.07)	0.89* (.06)	0.20 (.10)	2.00	0.87
Fourth	75.50* (2.99)	0.99* (.05)	0.67* (.05)	0.32* (.08)	4.00	0.92
Fifth	86.28* (3.06)	0.90* (.05)	0.67* (.06)	0.23 (.09)	2.50	0.54
Sixth	106.98* (4.40)	0.69* (.07)	0.73* (.06)	−0.04 (.10)	0.40	0.11

Note. SE = standard error.
[a]The estimated coefficients of the second slope from an added growth model.
[b]Standardized effect sizes = $\hat{\beta}_2$/SE, which is a strategy developed by Satorra and Saris (1985), was used to calculate power.
*$p < .001$.

greater than the second slope (spring) in all grades except for sixth; however, those differences were not statistically significant except in the fourth-grade condition (Table 3). Moreover, the pattern of reduced seasonal effects for progressively higher grades did not generalize to the sample of special education students. Instead, the magnitude was greatest in fourth grade (difference in slopes = 0.32, $p < .001$), somewhat smaller in third (0.20, $p > .001$) and fifth (0.23, $p > .001$), and hardly perceptible in second (0.09, $p > .001$) and sixth grades (−0.04, $p < .001$). The range of standardized effect sizes and power was variable across grades. Based on the results, it appears that, in general, the observed slopes in the sample of special education students were relatively consistent across seasons—in all conditions but fourth grade.

Discussion

Many of the published estimates of CBM-R annual growth rates rely on the assumption that the pattern is consistent and linear throughout the academic year (Deno et al., 2001; Fuchs et al., 1993). Although there is evidence for seasonal effects within some studies (Ardoin & Christ, 2009; Fuchs et al., 1993; Graney et al., 2009), the phenomenon of seasonal effects are generally ignored within the professional literature. This study yielded (a) estimates for the magnitude CBM-R annual growth along with those within the fall and spring seasons, and estimates were derived for both general education and special education populations; (b) both linear and piecewise growth models; and (c) hypothesis tests to examine whether fall growth is typically more substantial than spring growth.

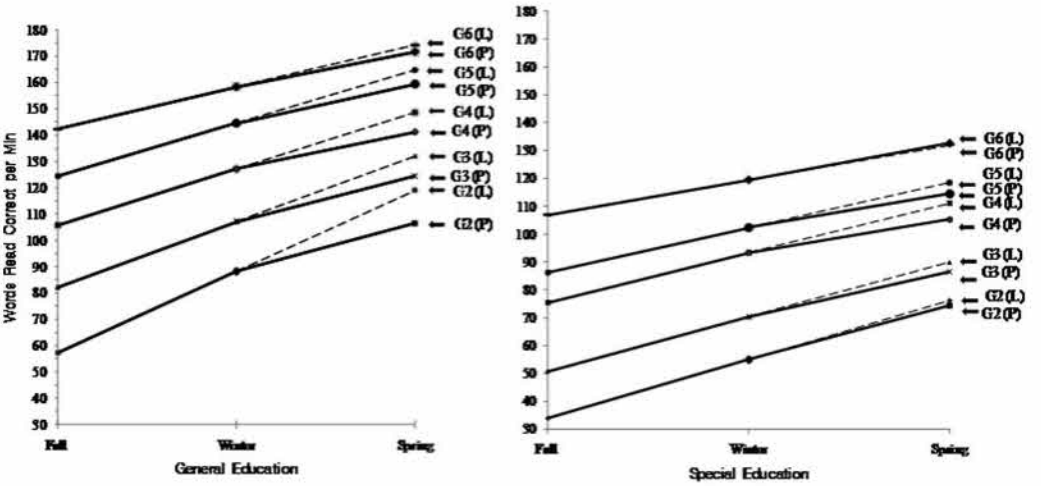


Figure 1. Differences for linear (L, solid line) and piecewise (P, dashed line) models of growth (G) for students in second through sixth grade within either the general (left) or special education population (right)

CBM-R growth was modeled and analyzed for five primary grades (second through sixth) and two student populations within and across seasons.

The results of this study support four general conclusions. Descriptive data make clear that, first, annual growth is greater among students in the general education population than for those in the special education population, especially in the fall season; and, second, more growth occurs in the early grades with less in the upper grades. Modeling and inferential analysis provide support for the third and fourth conclusions that follow. That is (third conclusion), there is a significant *seasonal effect* such that more growth occurs in the fall season than in the spring season for general education, but the magnitude of that effect declines with each progressive grade level. Finally, the seasonal effect is less pronounced among students in the special education population. In general, linear growth should not be assumed, seasonal effects should be expected—at least within the general education population—and piecewise models tend to fit better than linear models of growth.

Grade and Population Effects

Larger magnitudes of growth were observed for students in general education as compared to special education. This outcome was observed across five districts where CBM-R was used to screen, monitor, and evaluate instructional effects as part of a problem solving model. The systematic and ongoing use of CBM-R for these purposes did not establish equivalent growth rates between the student populations. Despite inconsistent growth rates across student groups, the rate of growth among students in the special education population was greater than observed in other studies (Deno et al., 2001; Fuchs et al., 1993). For example, the most recent estimates for average growth among students in the special education population were ~0.60 WRCM/week (Deno et al., 2001). That rate of growth was relatively consistent across grades. In contrast, the average growth estimates among students in the sample of students in special education used in this study were within the range of 0.67–1.17 WRCM/week (Table 3). Although cross-study comparisons cannot be used to establish that differences in systems-level practices (e.g., problem solving) and instructional practices were causally related to

the observed differences in growth rates, future research should investigate the phenomena more closely.

The results of this study converge with previous research (Deno et al., 2001; Fuchs et al., 1993) to support the conclusion that there is more CBM-R growth within the early primary grades. This phenomenon was consistent for both fall and spring seasons in both the general education and special education populations (Table 1). There was a steady decline for both general education and special education in weekly growth as grade level was increased. There was a corresponding decline in the magnitude of seasonal effects as grade levels increased. Results also support the conclusion that seasonal effects are more robust for students served within general education and less robust for students served within special education.

Seasonal Effect

The primary purpose of this research was to determine whether the rate of CBM-R growth is consistent across the year. Notwithstanding limitations, a seasonal effect was sufficiently robust and consistent across grades and student populations to warrant serious consideration. Estimates of weekly CBM-R growth within season are reported in Table 1. They are also depicted in Figure 1. Both visual analysis and statistical analysis of growth models support the conclusion that there was a *difference in the average rate of growth across seasons*, and that difference occurred in all grades, but was generally more robust in the lower grades of general education (Figure 1). With few exceptions, more growth occurred in the fall than in the spring, which is depicted as the differences in the slopes within Table 3.

The *magnitudes of seasonal effects were inconsistent* across grades and student populations. In general education, the difference between fall and spring slopes ranged from 0.14 to 0.69 per week, and all the differences were statistically significant with substantial effect sizes (Table 3). That effect is illustrated in Figure 1, which reveals that the interactions between seasonal effect and student population were most distinct in second grade. That is, the seasonal effect was most robust for the sample of students in second-grade general education classes, with a difference in growth from fall to winter of 0.69, and minor in comparison for students in second-grade special education classes, with a difference of 0.09.

The results of this study support the conclusion that growth differs between fall and spring seasons, which is inconsistent with a number of currently held views in the field. The consistency of the results across grades for general education begs the question of why this effect may be occur and what are the potential implications. There are a number of alternate hypotheses that offer some insight on the phenomenon and directions for future research.

Explanatory Hypotheses

This study was designed to examine *whether* seasonal effects are likely to occur in cases of CBM-R tri-annual assessment and challenge the assumption of linear growth. Indeed, the results do provide support for seasonal effects; however, this study was not designed to examine the cause of seasonal effects. That is left to future research to isolate the cause(s) and potential remedies to accelerate reading achievement within seasons with deficit growth. The result of both this study and Ardoin and Christ (2008) converge to provide evidence that the seasonal effect is likely to coincide with deficit levels of growth in the latter part of the academic year, or the spring season, especially for the students in the early primary general education population. It seems that changes in instructional conditions, or student motivation, might explain deficit rates of growth within the spring season unless there is an active intervention to prevent those changes. It seems that data-based decision making could function to reduce, or reverse, the pattern of seasonal effects that were replicated across this study and that of Ardoin and Christ.

Teacher expectations. There are a number of ways in which teacher expectations can have an influence on student behavior,

which might be conveyed through discrepant use of praise and punishment statements, content-relevant feedback, individualized attention, differentiated instruction, and variability in the opportunities to respond and engage in instructionally relevant practice (Good, 1981). Prior research yielded conclusions that there are indeed sometimes differences in teacher treatment of high- and low-achieving students over the course of the year, which affected the amount and type of instruction and feedback that students received at different points in the academic year (Good et al., 1980). Researchers concluded that teachers are more invested in influencing and shaping student behavior at the beginning of the school year than later in the year. Furthermore, the differentiation in teacher behavior may have sustained effects and become more influential as the school year progresses, resulting in lower achievement and growth later in the school year in low-achieving students. Other research also demonstrated the potential influence of teacher expectations, including expectation biases across ethnic groups (McKown & Weinstein, 2007) and ability groups (Rosenthal & Jacobson, 1966).

Classroom management. There is evidence that classroom management and organizational strategies are implemented with greater fidelity and effort in the early part of the academic year (Cameron, Connor, & Morrison, 2004). The same research provides evidence that that those educators who approach classroom management and organization with the greatest fervor early in the year are those same educators whose practices drop off most substantially later in the academic year. Improved levels of engagement and instructional gain are clearly associated with good classroom management and organization, and are among better predictors of student achievement (Brophy, 1987). Taken together, research provides a foundation to postulate that differential levels of classroom management and organization across seasons might influence student achievement. This effect is hypothesized to be the greatest at lower grade levels, as younger students require greater direction in organization practices (Cameron et al., 2004). The pattern of seasonal effects observed in this study coincides with a classroom management and organizational hypothesis.

Intervention to accelerate seasonal growth. The results of this study indicate that the seasonal effect with deficit growth in the spring season was most pronounced in the middle primary grades (second and third) and were minimized for the special education population. The results of Graney et al. (2009) indicate that the seasonal effect can function so that there is deficit growth in the fall season and accelerated growth in the spring. As noted in the Introduction, the data set that was analyzed by Graney et al. (2009) derived from a university and school partnership to implement a data-based decision making and evidence-based interventions. Those factors might have influenced the patterns of seasonal effects. There is a substantial literature base to support the conclusion that data-based decision making can function to accelerate academic growth (Carnine & Granzin, 2001; Fuchs & Fuchs, 1986; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Fuchs, Fuchs, Hamlett, & Whinnery, 1991; Graney et al., 2009), and there is substantial theoretical support for the use of evidence-based practices, which were both components of service delivery that were targeted by the partnership program. It might be that progress monitoring and data use accelerated learning and ameliorated the deficit growth associated with seasonal effects. Indeed, students who are served within special education have explicit goals and their progress is monitored. Although it is not possible to draw any final conclusions, it may well be that the lack of convergence in seasonal patterns between Graney et al. and other studies—along with minimal seasonal effects observed for the special education population—provides further evidence for the use of data-based decision making to improve growth during otherwise deficit periods.

Limitations and Future Research

The sample studied was entirely from five relatively ethnically homogeneous dis-

tricts in a single midwestern state. Future research should attempt to corroborate these findings across a more diverse sample of students and educational programs. Finally, the results of this study derive from data collected at three points in the academic year (fall, winter, spring). Future research should examine whether the same conclusions would result from the analysis of a more robust data set (e.g., weekly, monthly) and with more precise dates that define the data collection schedule. Graney et al. (2009) correctly criticize the methods of this study and that of Ardoin and Christ (2008) because the precise dates of data collection were not used in the analysis; rather, both studies rely on data that are centered by month and not day. Although the authors of this and the prior study believe the conclusions would be substantially similar with different centering methods or more precise dates, the criticism ought to be accounted for in future research. Future research should investigate the nature of growth across summative and formative measurement conditions. Research should also examine the effect on educational decisions when the potential for seasonal effects is ignored and consistent annual growth is assumed.

Implications for School-Based Decision Making

CBM-R was developed to guide routine instructional decisions and enhance instructional effects (Deno, 1986), and it has emerged as one of the primary methods of assessment to guide problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention models of service delivery (National Association of State Directors of Education, 2005). As described in the Introduction of this study, CBM-R is used to establish expectations for both the level and rate of reading achievement throughout the primary grades; therefore, the results of this study provide at least two clear implications. First, there is good reason to question the validity of linear growth estimates as criteria to guide progress monitoring (i.e., discrepant in rate) and dual-discrepancy decisions (i.e., discrepant in both level and rate; Fuchs, Fuchs, McMaster & Al Otaiba, 2003). Second, seasonal effects can be either admired as a phenomenon or targeted for intervention. These implications are addressed in what follows.

First, if CBM-R estimates of growth are used to evaluate instructional effects and response to intervention, then those who interpret data should be aware that the magnitude of growth might be less in the spring than in the fall. There are two weighty implications that derive from this study. First, a dual-discrepancy approach to special education eligibility, which compares the observed level and rate of achievement to criterion standards (Fuchs, Fuchs, McMaster & Al Otaiba, 2003), might function to overidentify students during deficit seasons of growth and underidentify students during accelerated seasons of growth. If the data follow the pattern observed in this study, then it is likely that the dual-discrepancy model might underidentify students in the fall season and overidentify students in the spring season. That consequence is counter to the values of early intervention and prevention inherent within problem solving and response to intervention. Researchers and practitioners should pay close attention to the patterns of growth and the potential for seasonal effects to influence decisions at the local level. Accelerated rates of growth should be expected in the fall.

Second, the pattern of seasonal effects that are observed across populations and studies provide preliminary evidence that goal setting, progress monitoring and data-based decision making can minimize, eliminate or reverse the pattern of seasonal effects (e.g., Graney et al., 2009). Although future research is necessary, it is a working thesis that the systematic use of data to evaluate instructional effects and select effective instructional strategies effectively accelerates learning and prevents problems at the individual, group, and systems level. After all, data-based decision making is a central tenet of problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention (National Association of State Directors of Education, 2005); however, data must be both systematically collected *and* sys-

tematically used to establish effects (Fuchs et al., 1992, 1991; Fuchs & Fuchs, 1986). The mere act of data collection without data use typically fails to establish effects (Fuchs & Fuchs, 1986). The explanatory hypotheses proposed in the previous section establish that deficit growth in the spring season might be remediated at the systems level if data are systematically collected and systematically used. If detected, this problem (seasonal effect) is ripe for problem analysis at the local level. Deficit growth within a single season should not be admired as an acceptable phenomenon, but, instead, it should be targeted for analysis and intervention.

Christ (2008) defined problem analysis as the "collection, summary, and use of information to systematically test, reject, or verify relevant hypotheses to establish problem solutions" (p. 159). That is, deficit rates of growth in the spring season might be remediated if expectations remain high and data are used to evaluate and maintain high standards for instructional effects. There is no reason to assume that the seasonal effect is inherent to the educational system. Instead, hypotheses related to the causal and maintaining variables should be developed—such as those presented earlier—and tested through intervention to remediate the problem. It is likely that the pattern of seasonal effects is causally related to what occurs in the classroom; and what occurs in the classroom can be influenced by the systematic used of data to evaluate of response to intervention for universal, supplemental, and intensive tiers of service delivery. This study provides evidence of seasonal effects as a potential target for intervention at the system, grade, or classroom level.

References

AIMSweb. (2002). *Reading curriculum-based measurement (R-CBM) standard oral reading fluency assessment passages.* Retrieved October 2003, from http://www.aimsweb.com

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum based measurement slope estimate using data from tri-annual universal screenings *School Psychology Review, 37,* 109–125.

Bollman, K. A., Silberglitt, B., & Gibbons, K. A. (2007). The SCRED Model: Incorporating systems-level organization and a multi-tiered problem-solving process for intervention delivery. In M. Burns, S. Jimerson, & A. VanDerHayden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 321–332). New York: Springer.

Brophy, J. (1987). Synthesis of research on strategies for motivating students to learn. *Educational Leadership, 45*(2), 40–48.

Cameron, C. E., Connor, C. M., & Morrison, F. J. (2004). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology, 43,* 61–85.

Carnine, D., & Granzin, A. (2001). Setting learning expectations for students with disabilities. *School Psychology Review, 30,* 466–472.

Christ, T. J. (2008). Best practices in problem analysis. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 159–176). Bethesda, MD: National Association of School Psychologists.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74,* 68.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219–232.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15,* 358–374.

Deno, S. L. (2002). Problem solving as best practices. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 37–56). Bethesda, MD: National Association of School Psychologists.

Deno, S. L. (2005). Problem solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 10–42). New York: Guilford Press.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30,* 507–524.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). An introduction to latent variable growth curve modeling: Concepts, issues, and applications (2nd ed.). Mahwah, NY: Lawrence Erlbaum Associates.

Flora, D. B. (2008). Specifying piecewise latent trajectory models for longitudinal data. *Structural Equation Modeling, 15,* 513–533.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53,* 3–14.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10,* 43–52.

Fuchs, L. S., Fuchs, D., & Stecker, P. M. (1989). Effects of curriculum-based measurement on teachers' instructional planning. *Journal of Learning Disabilities, 22,* 51–59.

Fuchs, L. S., & Shinn, M. R. (1989). Writing CBM IEP objectives. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 131–152). New York: Guildford Press.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58,* 436–450.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22,* 27–48.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Whinnery, K. (1991). Effects of goal line feedback on level, slope, and stability of performance within curriculum-based measurement. *Learning Disabilities Research and Practice, 6,* 66–74.

Fuchs, D., Fuchs, L. S., McMaster, K. N., & Al Otaiba, S. (2003). Identifying children at risk for reading failure: Curriculum-based measurement and the dual-discrepancy approach. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 431–449). New York: Guilford Press.

Good, T. L. (1981). Teacher expectations and student perceptions: A decade of research. *Educational Leadership, 38*(5), 415–422.

Good, T., Cooper, H., & Blakey, X. (1980). Classroom interaction as a function of teacher expectations, student sex and time of year. *Journal of Educational Psychology, 72*, 378–385.

Graney, S. B., Missall, K. N., Martinez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School* Psychology, 47, 121–142.

Howe, K. B., & Shinn, M. M. (2002). Standard *reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Retrieved October 2004, from http://www.aimsweb.com

Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education for Students Placed at Risk, 9,* 169–188.

McKown, C., Weinstein, R. S. (2007). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*, 235–261.

National Association of State Directors of Education. (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: Author.

National Reading Panel. (2000). *Report of the National Reading Panel Subgroups: Teaching children to read* (No. NIH 00–4754). Bethesda, MD: National Institute of Child Health and Human Development.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports, 19,* 115–118.

Satorra, A., & Saris, W. E. (1985). The power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50*, 83–90.

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education, 34,* 164–172.

Shinn, M. R. (1988). Development of curriculum-based local norms for use in special education decision-making. *School Psychology Review, 17,* 61–80.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children.* New York: Guildford Press.

Shinn, M. (2008). Best practices for using curriculum based measurement in a problem solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 243–262). Bethesda, MD: National Association of School Psychologists.

Stage, S. A. (2001). Program evaluation using hierarchical linear modeling with curriculum-based measurement reading probes. *School Psychology Quarterly, 16,* 91–112.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41,* 85–120.

Date Received: May 30, 2007
Date Accepted: April 21, 2010
Action Editor: Tanya Eckert ■

Theodore J. Christ, PhD, is an associate professor in the Educational Psychology Department at the University of Minnesota in Minneapolis—Saint Paul. He is the codirector of the Research Institute for Problem Solving. He is engaged in research and training to evaluate, develop, and enhance the tools and procedures that enable data-based decision making and problem solving within school settings. Interests include progress monitoring, curriculum-based measurement, direct behavior ratings, and computer adaptive testing.

Benjamin Silberglitt completed his PhD in Educational Psychology at the University of Minnesota at Minneapolis—Saint Paul. His areas of interest include student assessment school organization and system-level change, problem solving, response to intervention, and data-based decision making. He is director of Software Applications for TIES, a technology cooperative of 39 Minnesota school districts.

Seungsoo Yeo, PhD, is a researcher at Korean Educational Development Institute. He received his PhD from the University of Minnesota at Twin Cities in 2008. His primary research interests include curriculum-based measurement, response to intervention, and learning disabilities. His methodological work has focused on latent growth modeling and multilevel meta-analysis.

Damien C. Cormier, MA, is a school psychology doctoral student in the Educational Psychology Department at the University of Minnesota. He is a graduate research assistant at the National Center on Educational Outcomes. His research interests are in the areas of large-scale assessment policy and practice, measurement issues in individual assessment, high-functioning autism, and cultural and linguistic biases in assessment.

Copyright of School Psychology Review is the property of National Association of School Psychologists and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

I was mistaken when I indicated that the Reading Fluency Monitor was developed by Read Naturally. It was actually developed by the RMC Research Corporation out of Portsmouth, New Hampshire.

Sorry for the misunderstanding,

- Dave

David Heistad, Ph.D., Executive Director
Research and Evaluation Department
Bloomington Public Schools - ISD 271
1350 W. 106th Street
Bloomington, MN 55431
ph: 952-681-6485
cell: 952-250-5971

Karla Ramy

From:Karla Ramy
Sent:24 Oct 2013 12:43:50 +0000
To:What Works
Subject:Read Naturally - Tom Ihnot
Attachments:Christ Silberglitt Yeo Cormier (2010).pdf, Heistad.docx
Dear Jill Constantine,

In an August 26, 2013 letter from What Works Clearinghouse (WWC), Jill Constantine stated:

> *Order of findings.* We appreciate and agree with your suggestion of July 17, in which you propose that the WWC list findings in order of their strength, consistent with the presentation on the web. Therefore, we will revise the order of findings listed in the text on the front page of the report, the associated table, and the website summary for the three reports with findings in more than one domain. We anticipate that the revised intervention reports will be posted to the web by the end of September. (Reference :QR2010013, WWC2304)

I look forward to this change being executed as soon as possible. Please email me a verification that the change in order of findings will read:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

Also, I received an October 18, 2013 email (Reference: QR2013002) responding to my concerns about the Heistad and Arvans studies. Thank you for your response. Concerning the Heistad study, you stated that the measure was over-aligned with the intervention as words used in the Read Naturally curriculum also appear in the Reading Fluency Monitor. I do not understand this reasoning. Please explain it to me in more detail. I am concerned that the following facts are not being properly evaluated:

- Read Naturally passages were developed by Read Naturally staff as curriculum.
- The Reading Fluency Monitor was developed by RMC Research Corporation, located in Portsmouth, New Hampshire, as a benchmark assessment tool.
- Read Naturally curriculum passages and the Reading Fluency Monitor are completely different tools developed by different companies.
- Dr. Heistad clearly states that he made a mistake in the test instruments section of his study. His verification of this is attached.

Including Dr. Heistad's statistically significant gains in oral reading fluency for third-grade students across multiple schools would change the effectiveness rating in the Beginning Reading section for Read Naturally from mixed to positive effects.

The Arvans study clearly shows that Read Naturally programs are research proven when it comes to oral reading fluency gains. The group using the Read Naturally strategy had a large effect size of .81. When you conduct a short eight-week study, it is an accepted practice to do an extrapolation for a school

year. (Christ – Attached) What Works Clearinghouse should consider changing its policy on extrapolations.

I look forward to a prompt listing and announcement of the new wording on Read Naturally's effectiveness rating for the Beginning Reading Section. Also, I would appreciate a reevaluation of the over-alignment issue with the Heistad study and further consideration of the Arvans study.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:25 Oct 2013 02:20:09 +0000
To:'kramy@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot WWCPC 4299
Dear Mr. Ihnot,

We have forwarded your email to the quality review team. They are drafting a response to your questions.

Thank you,

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Thursday, October 24, 2013 8:44 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Dear Jill Constantine,

In an August 26, 2013 letter from What Works Clearinghouse (WWC), Jill Constantine stated:

> *Order of findings.* We appreciate and agree with your suggestion of July 17, in which you propose that the WWC list findings in order of their strength, consistent with the presentation on the web. Therefore, we will revise the order of findings listed in the text on the front page of the report, the associated table, and the website summary for the three reports with findings in more than one domain. We anticipate that the revised intervention reports will be posted to the web by the end of September. (Reference :QR2010013, WWC2304)

I look forward to this change being executed as soon as possible. Please email me a verification that the change in order of findings will read:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

Also, I received an October 18, 2013 email (Reference: QR2013002) responding to my concerns about the Heistad and Arvans studies. Thank you for your response. Concerning the Heistad study, you stated that the measure was over-aligned with the intervention as words used in the Read Naturally curriculum also appear in the Reading Fluency Monitor. I do not understand this reasoning. Please explain it to me in more detail. I am concerned that the following facts are not being properly evaluated:

- Read Naturally passages were developed by Read Naturally staff as curriculum.

- The Reading Fluency Monitor was developed by RMC Research Corporation, located in Portsmouth, New Hampshire, as a benchmark assessment tool.
- Read Naturally curriculum passages and the Reading Fluency Monitor are completely different tools developed by different companies.
- Dr. Heistad clearly states that he made a mistake in the test instruments section of his study. His verification of this is attached.

Including Dr. Heistad's statistically significant gains in oral reading fluency for third-grade students across multiple schools would change the effectiveness rating in the Beginning Reading section for Read Naturally from mixed to positive effects.

The Arvans study clearly shows that Read Naturally programs are research proven when it comes to oral reading fluency gains. The group using the Read Naturally strategy had a large effect size of .81. When you conduct a short eight-week study, it is an accepted practice to do an extrapolation for a school year. (Christ – Attached) What Works Clearinghouse should consider changing its policy on extrapolations.

I look forward to a prompt listing and announcement of the new wording on Read Naturally's effectiveness rating for the Beginning Reading Section. Also, I would appreciate a reevaluation of the over-alignment issue with the Heistad study and further consideration of the Arvans study.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally® CELEBRATING 20 YEARS

What Works

From:What Works
Sent:28 Oct 2013 16:31:05 +0000
To:'kramy@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot WWCPC 4299
Dear Mr. Ihnot,

The quality review team reviewed your latest correspondence. All four Read Naturally intervention reports (Adolescent Literacy, Beginning Reading, English Language Learners, and Students with Learning Disabilities) will be updated on the WWC website by November 1. Regarding the Reading Fluency Monitor, overalignment is an issue for a measure when students in the treatment group have previously seen material that is on the test and students in the comparison group have not. It is not based on who developed the intervention or the measure. Finally, the WWC's current policies do not support extrapolation of findings. Please let us know if you have other concerns.

Thank you,

What Works Clearinghouse

From: WhatWorks
Sent: Thursday, October 24, 2013 10:20 PM
To: 'kramy@readnaturally.com'
Subject: RE: Read Naturally - Tom Ihnot WWCPC 4299

Dear Mr. Ihnot,

We have forwarded your email to the quality review team. They are drafting a response to your questions.

Thank you,

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Thursday, October 24, 2013 8:44 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Dear Jill Constantine,

In an August 26, 2013 letter from What Works Clearinghouse (WWC), Jill Constantine stated:

> *Order of findings.* We appreciate and agree with your suggestion of July 17, in which you propose that the WWC list findings in order of their strength, consistent with the presentation on the web. Therefore, we will revise the order of findings listed in the text on the front page of the report, the associated table, and the website summary for the three

reports with findings in more than one domain. We anticipate that the revised intervention reports will be posted to the web by the end of September. (Reference :QR2010013, WWC2304)

I look forward to this change being executed as soon as possible. Please email me a verification that the change in order of findings will read:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

Also, I received an October 18, 2013 email (Reference: QR2013002) responding to my concerns about the Heistad and Arvans studies. Thank you for your response. Concerning the Heistad study, you stated that the measure was over-aligned with the intervention as words used in the Read Naturally curriculum also appear in the Reading Fluency Monitor. I do not understand this reasoning. Please explain it to me in more detail. I am concerned that the following facts are not being properly evaluated:

- Read Naturally passages were developed by Read Naturally staff as curriculum.
- The Reading Fluency Monitor was developed by RMC Research Corporation, located in Portsmouth, New Hampshire, as a benchmark assessment tool.
- Read Naturally curriculum passages and the Reading Fluency Monitor are completely different tools developed by different companies.
- Dr. Heistad clearly states that he made a mistake in the test instruments section of his study. His verification of this is attached.

Including Dr. Heistad's statistically significant gains in oral reading fluency for third-grade students across multiple schools would change the effectiveness rating in the Beginning Reading section for Read Naturally from mixed to positive effects.

The Arvans study clearly shows that Read Naturally programs are research proven when it comes to oral reading fluency gains. The group using the Read Naturally strategy had a large effect size of .81. When you conduct a short eight-week study, it is an accepted practice to do an extrapolation for a school year. (Christ – Attached) What Works Clearinghouse should consider changing its policy on extrapolations.

I look forward to a prompt listing and announcement of the new wording on Read Naturally's effectiveness rating for the Beginning Reading Section. Also, I would appreciate a reevaluation of the over-alignment issue with the Heistad study and further consideration of the Arvans study.

Sincerely,

Tom Ihnot
CEO

2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Karla Ramy

From:Karla Ramy
Sent:29 Oct 2013 20:30:56 +0000
To:What Works
Subject:To Jill Constantine - Heistad Study
Dear Jill Constantine,

Dr. Dave Heistad, author of the Heistad study on the Read Naturally intervention strategy, does not think the Reading Fluency Monitor Benchmark Assessment was over aligned with the Read Naturally intervention.

Dr. Heistad sent me the following email:

From: Dave Heistad [mailto:dheistad@bloomington.k12.mn.us]
Sent: Monday, October 28, 2013 3:00 PM
To: info@readnaturally.com
Subject: Heistad study

"We do not feel that the Read Naturally Intervention was "over-aligned" with the measurement system. Students in the Read Naturally group did not see material on the assessment previous to being assessed as part of the study. The measures were purchased from a company that developed the passages independent of the Read Naturally intervention strategies.

Vocabulary in the intervention strategies are no more likely to occur in the Reading Fluency Monitor passages than, let's say, DIBELS passages or Aimsweb passages."

- Dave

There must be some misunderstanding. Reading Fluency Monitor and the Read Naturally strategy are not aligned in anyway.
Let me know if I should send you the third grade Reading Fluency Monitor guidelines and passages and Read Naturally intervention passages for levels 2.0, 2.5, 3.0, 3.5, 4.0 and 4.5. If you need any additional information, please contact me. Also, feel free to contact Dr. Heistad.
I would appreciate it if you could provide a reevaluation.

Sincerely,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally®
CELEBRATING
20 YEARS

Karla Ramy

From:Karla Ramy
Sent:29 Oct 2013 20:29:43 +0000
To:What Works
Subject:To Neil Seftor - Heistad Study
Dear Neil Seftor,

Dr. Dave Heistad, author of the Heistad study on the Read Naturally intervention strategy, does not think the Reading Fluency Monitor Benchmark Assessment was over aligned with the Read Naturally intervention.

Dr. Heistad sent me the following email:

From: Dave Heistad [mailto:dheistad@bloomington.k12.mn.us]
Sent: Monday, October 28, 2013 3:00 PM
To: info@readnaturally.com
Subject: Heistad study

"We do not feel that the Read Naturally Intervention was "over-aligned" with the measurement system. Students in the Read Naturally group did not see material on the assessment previous to being assessed as part of the study. The measures were purchased from a company that developed the passages independent of the Read Naturally intervention strategies.

Vocabulary in the intervention strategies are no more likely to occur in the Reading Fluency Monitor passages than, let's say, DIBELS passages or Aimsweb passages."

- Dave

There must be some misunderstanding. Reading Fluency Monitor and the Read Naturally strategy are not aligned in anyway.
Let me know if I should send you the third grade Reading Fluency Monitor guidelines and passages and Read Naturally intervention passages for levels 2.0, 2.5, 3.0, 3.5, 4.0 and 4.5. If you need any additional information, please contact me. Also, feel free to contact Dr. Heistad.
I would appreciate it if you could provide a reevaluation.

Sincerely,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally®
CELEBRATING
20 YEARS

What Works

From:What Works
Sent:1 Nov 2013 16:09:15 +0000
To:'kramy@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot WWCPC 4299
Dear Mr. Ihnot,

All four Read Naturally intervention reports (Adolescent Literacy, Beginning Reading, English Language Learners, and Students with Learning Disabilities) have been updated on the WWC website. You can access the updated reports via the links below. The quality review team is currently considering your email from 10/30/2013 and preparing a response.

- Adolescent Literacy: http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=625
- Beginning Reading: http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=407
- English Language Learners: http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=408
- Students with Learning Disabilities: http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=409

Thank you,

What Works Clearinghouse

From: What Works
Sent: Monday, October 28, 2013 12:31 PM
To: 'kramy@readnaturally.com'
Subject: RE: Read Naturally - Tom Ihnot WWCPC 4299

Dear Mr. Ihnot,

The quality review team reviewed your latest correspondence. All four Read Naturally intervention reports (Adolescent Literacy, Beginning Reading, English Language Learners, and Students with Learning Disabilities) will be updated on the WWC website by November 1. Regarding the Reading Fluency Monitor, overalignment is an issue for a measure when students in the treatment group have previously seen material that is on the test and students in the comparison group have not. It is not based on who developed the intervention or the measure. Finally, the WWC's current policies do not support extrapolation of findings. Please let us know if you have other concerns.

Thank you,

What Works Clearinghouse

From: WhatWorks
Sent: Thursday, October 24, 2013 10:20 PM
To: 'kramy@readnaturally.com'
Subject: RE: Read Naturally - Tom Ihnot WWCPC 4299

Dear Mr. Ihnot,

We have forwarded your email to the quality review team. They are drafting a response to your questions.

Thank you,

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Thursday, October 24, 2013 8:44 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Dear Jill Constantine,

In an August 26, 2013 letter from What Works Clearinghouse (WWC), Jill Constantine stated:

> *Order of findings.* We appreciate and agree with your suggestion of July 17, in which you propose that the WWC list findings in order of their strength, consistent with the presentation on the web. Therefore, we will revise the order of findings listed in the text on the front page of the report, the associated table, and the website summary for the three reports with findings in more than one domain. We anticipate that the revised intervention reports will be posted to the web by the end of September. (Reference :QR2010013, WWC2304)

I look forward to this change being executed as soon as possible. Please email me a verification that the change in order of findings will read:

> Read Naturally® programs were found to have potentially positive effects on general reading achievement, mixed effects on reading fluency, and no discernible effects on alphabetics and comprehension for beginning readers.

Also, I received an October 18, 2013 email (Reference: QR2013002) responding to my concerns about the Heistad and Arvans studies. Thank you for your response. Concerning the Heistad study, you stated that the measure was over-aligned with the intervention as words used in the Read Naturally curriculum also appear in the Reading Fluency Monitor. I do not understand this reasoning. Please explain it to me in more detail. I am concerned that the following facts are not being properly evaluated:

- Read Naturally passages were developed by Read Naturally staff as curriculum.
- The Reading Fluency Monitor was developed by RMC Research Corporation, located in Portsmouth, New Hampshire, as a benchmark assessment tool.
- Read Naturally curriculum passages and the Reading Fluency Monitor are completely different tools developed by different companies.

- Dr. Heistad clearly states that he made a mistake in the test instruments section of his study. His verification of this is attached.

Including Dr. Heistad's statistically significant gains in oral reading fluency for third-grade students across multiple schools would change the effectiveness rating in the Beginning Reading section for Read Naturally from mixed to positive effects.

The Arvans study clearly shows that Read Naturally programs are research proven when it comes to oral reading fluency gains. The group using the Read Naturally strategy had a large effect size of .81. When you conduct a short eight-week study, it is an accepted practice to do an extrapolation for a school year. (Christ – Attached) What Works Clearinghouse should consider changing its policy on extrapolations.

I look forward to a prompt listing and announcement of the new wording on Read Naturally's effectiveness rating for the Beginning Reading Section. Also, I would appreciate a reevaluation of the over-alignment issue with the Heistad study and further consideration of the Arvans study.

Sincerely,

Tom Ihnot
CEO

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally® CELEBRATING 20 YEARS

What Works

From:What Works
Sent:13 Nov 2013 17:34:58 +0000
To:'Karla Ramy';'info@readnaturally.com'
Subject:RE: To Jill Constantine - Heistad Study

Dear Mr. Ihnot,

Thank you for this additional information regarding the Heistad study. Study authors may contact the WWC directly at any time to provide information on their study. If Dr. Heistad has additional or new information to provide, he may send it through the Help Desk.

Thank you,

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Tuesday, October 29, 2013 4:31 PM
To: What Works
Subject: To Jill Constantine - Heistad Study

Dear Jill Constantine,

Dr. Dave Heistad, author of the Heistad study on the Read Naturally intervention strategy, does not think the Reading Fluency Monitor Benchmark Assessment was over aligned with the Read Naturally intervention.

Dr. Heistad sent me the following email:

From: Dave Heistad [mailto:dheistad@bloomington.k12.mn.us]
Sent: Monday, October 28, 2013 3:00 PM
To: info@readnaturally.com
Subject: Heistad study

"We do not feel that the Read Naturally Intervention was "over-aligned" with the measurement system. Students in the Read Naturally group did not see material on the assessment previous to being assessed as part of the study. The measures were purchased from a company that developed the passages independent of the Read Naturally intervention strategies.

Vocabulary in the intervention strategies are no more likely to occur in the Reading Fluency Monitor passages than, let's say, DIBELS passages or Aimsweb passages."

- Dave

There must be some misunderstanding. Reading Fluency Monitor and the Read Naturally strategy are not aligned in anyway.
Let me know if I should send you the third grade Reading Fluency Monitor guidelines and passages and Read Naturally intervention passages for levels 2.0, 2.5, 3.0, 3.5, 4.0 and 4.5. If you need any additional information, please contact me. Also, feel free to contact Dr. Heistad.
I would appreciate it if you could provide a reevaluation.

Sincerely,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Karla Ramy

From:Karla Ramy
Sent:16 Dec 2013 15:35:11 +0000
To:What Works
Subject:To Jill Constantine - from Tom Ihnot

Jill Constantine,

Thank you for your confirmation of Dr. Heistad's email of 12/10/13. In the email he verifies that he made a mistake in the test instrument section of his report, *The Effects of Read Naturally on Grade 3 Reading: A Study in the Minneapolis Public Schools*. Dr. Heistad states in his email, "I was mistaken in saying the Reading Fluency Monitor was 'developed' by Read Naturally Inc. What I should have said is that the passages for the Reading Fluency Monitor were developed by the RMC Research Corporation out of Portsmouth, New Hampshire and later purchased by Read Naturally. I do not feel that the Read Naturally Intervention was 'over-aligned' with the measurement system. I believe that the vocabulary present in the intervention strategies are no more likely to occur in the Reading Fluency Monitor passages than, let's say, DIBELS passages or AIMSWEB passages."

The following facts are clear:

- Read Naturally passages were developed by Read Naturally staff as curriculum.
- The Reading Fluency Monitor was developed by RMC Research Corporation, located in Portsmouth, New Hampshire, as a benchmark assessment tool.
- Read Naturally curriculum passages and the Reading Fluency Monitor are completely different tools developed by different companies.

Based on this information, WWC should include Dr. Heistad's statistically significant gains in oral reading fluency for third-grade students across multiple schools in its report for Read Naturally. Including this data should change the effectiveness rating in the Beginning Reading section for Read Naturally from mixed to positive effects.

I look forward to a positive decision from WWC's reevaluation of the Heistad study based on the new information Dr. Heistad has provided.

Thank you,
Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally®
CELEBRATING
20 YEARS

WhatWorks

From:WhatWorks
Sent:17 Dec 2013 05:49:41 +0000
To:'kramy@readnaturally.com'
Subject:RE: To Jill Constantine - from Tom Ihnot

Dear Mr. Ihnot,

Thank you for your email. The WWC welcomes updated information from study authors. A previously reviewed study can be updated to reflect new information received after the initial review. If an updated review of a study changes the findings of an intervention report, the report will be updated to reflect the new information.

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Monday, December 16, 2013 10:35 AM
To: What Works
Subject: To Jill Constantine - from Tom Ihnot

Jill Constantine,

Thank you for your confirmation of Dr. Heistad's email of 12/10/13. In the email he verifies that he made a mistake in the test instrument section of his report, *The Effects of Read Naturally on Grade 3 Reading: A Study in the Minneapolis Public Schools*. Dr. Heistad states in his email, "I was mistaken in saying the Reading Fluency Monitor was 'developed' by Read Naturally Inc. What I should have said is that the passages for the Reading Fluency Monitor were developed by the RMC Research Corporation out of Portsmouth, New Hampshire and later purchased by Read Naturally. I do not feel that the Read Naturally Intervention was 'over-aligned' with the measurement system. I believe that the vocabulary present in the intervention strategies are no more likely to occur in the Reading Fluency Monitor passages than, let's say, DIBELS passages or AIMSWEB passages."

The following facts are clear:

- Read Naturally passages were developed by Read Naturally staff as curriculum.
- The Reading Fluency Monitor was developed by RMC Research Corporation, located in Portsmouth, New Hampshire, as a benchmark assessment tool.
- Read Naturally curriculum passages and the Reading Fluency Monitor are completely different tools developed by different companies.

Based on this information, WWC should include Dr. Heistad's statistically significant gains in oral reading fluency for third-grade students across multiple schools in its report for Read Naturally. Including this data should change the effectiveness rating in the Beginning Reading section for Read Naturally from mixed to positive effects.

I look forward to a positive decision from WWC's reevaluation of the Heistad study based on the new information Dr. Heistad has provided.

Thank you,
Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

Karla Ramy

From:Karla Ramy
Sent:6 Feb 2014 15:47:54 +0000
To:What Works
Subject:Read Naturally - Tom Ihnot
Susanne James-Burdumy,

I understand that you have corresponded with Dr. Dave Heistad concerning his study, The Effects of Read Naturally on Grade 3 Reading: A Study in the Minneapolis Public Schools.

According to Dr. Heistad, he confirmed that the assessment passages used in his study were independent from the Read Naturally curriculum passages. In addition, he responded to other questions you raised.

Thank you for reconsidering Dr. Heistad's study.

Please let me know when you will make a decision on including Dr. Heistad's study. His study showed statistically significant gains in oral reading fluency for third grade students across multiple schools. I am hopeful the Heistad Grade 3 Reading fluency results will be included in the WWC report on Read Naturally.

Sincerely,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

WhatWorks

From:WhatWorks
Sent:7 Feb 2014 15:57:10 +0000
To:'kramy@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot (WWC 4299)
Dear Mr. Ihnot,

Thank you for your email. We will let you know when the report is updated.

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Thursday, February 06, 2014 10:48 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Susanne James-Burdumy,

I understand that you have corresponded with Dr. Dave Heistad concerning his study, The Effects of Read Naturally on Grade 3 Reading: A Study in the Minneapolis Public Schools.

According to Dr. Heistad, he confirmed that the assessment passages used in his study were independent from the Read Naturally curriculum passages. In addition, he responded to other questions you raised.

Thank you for reconsidering Dr. Heistad's study.

Please let me know when you will make a decision on including Dr. Heistad's study. His study showed statistically significant gains in oral reading fluency for third grade students across multiple schools. I am hopeful the Heistad Grade 3 Reading fluency results will be included in the WWC report on Read Naturally.

Sincerely,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax

ReadNaturally®
CELEBRATING
20 YEARS

WhatWorks

From:WhatWorks
Sent:27 Feb 2014 17:30:11 +0000
To:'kramy@readnaturally.com'
Subject:RE: Read Naturally - Tom Ihnot (WWC 4299)
Hello Mr. Ihnot,

The revised Read Naturally report is now live on the WWC website and is available at: http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=407.

Thank you,

What Works Clearinghouse

From: WhatWorks
Sent: Friday, February 07, 2014 9:57 AM
To: 'kramy@readnaturally.com'
Subject: RE: Read Naturally - Tom Ihnot (WWC 4299)

Dear Mr. Ihnot,

Thank you for your email. We will let you know when the report is updated.

What Works Clearinghouse

From: Karla Ramy [mailto:kramy@readnaturally.com]
Sent: Thursday, February 06, 2014 10:48 AM
To: What Works
Subject: Read Naturally - Tom Ihnot

Susanne James-Burdumy,

I understand that you have corresponded with Dr. Dave Heistad concerning his study, The Effects of Read Naturally on Grade 3 Reading: A Study in the Minneapolis Public Schools.

According to Dr. Heistad, he confirmed that the assessment passages used in his study were independent from the Read Naturally curriculum passages. In addition, he responded to other questions you raised.

Thank you for reconsidering Dr. Heistad's study.

Please let me know when you will make a decision on including Dr. Heistad's study. His study showed statistically significant gains in oral reading fluency for third grade students across multiple schools. I am hopeful the Heistad Grade 3 Reading fluency results will be included in the WWC report on Read Naturally.

Sincerely,

Tom Ihnot

Read Naturally, Inc.
2945 Lone Oak Drive, Suite 190
St. Paul, MN 55121
www.readnaturally.com
www.oneminutereader.com
651-452-4085
800-788-4085
651-452-9204 - fax