The handle http://hdl.handle.net/1887/67118 holds various files of this Leiden University dissertation.

Author: Bosch, R.M. van den
Title: Picturing student progress : teachers' comprehension of curriculum-based measurement progress graphs
Issue Date: 2018-11-22
IMPROVING TEACHERS’ COMPREHENSION OF CURRICULUM-BASED MEASUREMENT PROGRESS GRAPHS

Submitted as:
Chapter 5

ABSTRACT

We examined three instructional approaches for improving teachers’ CBM graph comprehension, each differing in the extent to which reading the data, interpreting the data, and linking the data to instruction were emphasized. Participants were 164 elementary-school teachers who were randomly assigned to one of three CBM instructional approaches or a control condition. Instruction was delivered via videos. Prior to and after receiving instruction, teachers completed a CBM graph-comprehension task. They also evaluated the instructional videos. Teachers in the three instructional groups improved more in CBM graph comprehension than control teachers. Improvements were seen primarily in interpreting and linking the data to instruction, two important but difficult aspects of CBM graph comprehension. For interpreting the data, differences between the instructional groups were found. Teachers evaluated the videos positively. Results imply that teachers’ CBM graph comprehension can be improved via video instruction. Implications for teaching teachers to implement CBM are discussed.
IMPROVING TEACHERS’ COMPREHENSION OF CURRICULUM-BASED MEASUREMENT PROGRESS GRAPHS

Curriculum-based measurement (CBM) is a tool that enables teachers to closely monitor student reading progress and evaluate the effectiveness of instruction for students who struggle with reading (Deno, 1985). When used within Tier 2 and Tier 3 settings, student performance is measured frequently with short tasks and scores are placed on graphs to display student progress over time (see Figure 5.1, for a sample graph). Special and/or general education teachers view the graphs, often within the context of team meetings, to evaluate student progress and to determine whether there is a need for an instructional or a goal change, depending on whether progress is less than or greater than expected. When teachers respond to the data with instructional/goal changes, student performance improves; however, teachers often do not respond to the data, at least not without support (Stecker, Fuchs, & Fuchs, 2005).

Figure 5.1. Sample of a CBM progress graph.

In the late 1980s and early 1990s, L. S. Fuchs, Fuchs and colleagues investigated the effects of decision-making supports, delivered via computer-software, on teachers’ CBM implementation (see L. S. Fuchs & Fuchs, 1989; L. S. Fuchs & Fuchs, 2002; Stecker et al., 2005, for reviews). These supports assisted teachers in collecting, scoring, and graphing the data, and prompted teachers to change instruction or to raise the goal. In
later stages of development, the computer supports provided diagnostic skills-analysis and expert-feedback to guide teachers’ decisions about what to change and how to change instruction. Use of the computer supports led to improvements in teachers’ implementation of CBM decision-rules, selection of appropriate instructional changes, and design of diverse educational programs, which in turn led to improvements in student performance (L. S. Fuchs & Fuchs, 1989; L. S. Fuchs & Fuchs, 2002; Stecker et al., 2005).

However, even with computer supports, the teacher remained an important element in the decision-making process. For example, teachers were more accurate in timing instructional changes if they first formulated instructional decisions than if they received computer feedback immediately (L. S. Fuchs, Fuchs, & Hamlett, 1989a). L. S. Fuchs and Fuchs (1989) noted that sole reliance on computer supports might serve to distance teachers from the data, reducing effective data-based decision-making.

Despite the important role of the teacher in CBM data-based decision-making, researchers have only recently begun to examine teachers’ CBM data-based decision-making processes. The focus of the research to date has been on the first step in the decision-making process: Reading and interpreting the CBM progress graphs.

**Teachers’ Comprehension of CBM Progress Graphs**

The ability to read and interpret graphs is often referred to as graph comprehension (Friel, Curcio, & Bright, 2001). Friel et al. (2001) described three levels of graph comprehension: *reading the data* (i.e., extracting data from the graph), *reading between the data* (i.e., integrating and interpreting the graphed data), and *reading beyond the data* (i.e., evaluating and interpreting data within a given context).

Van den Bosch, Espin, Chung, and Saab (2017) applied Friel et al.’s (2001) framework to the study of CBM graph comprehension, referring to the three levels as *reading the data* (i.e., describing the CBM data as they appear on the graph), *interpreting the data* (i.e., integrating and interpreting relations between graph elements such as slope and goal), and *linking the data to instruction* (i.e., evaluating and interpreting the data within the instructional context). Building upon two earlier studies (Espin, Wayman, Deno, McMaster, & de Rooij, 2017; Wagner, Hammerschmidt-Sniderach, Espin, Seifert, & McMaster, 2017), van den Bosch et al. (2017) employed a think-aloud method to examine teachers’ CBM graph comprehension, and to compare it to the CBM graph comprehension of three small groups of “gold-standard” graph-reading experts: General graph-reading experts, education graph-reading experts, and CBM graph-reading experts. (The CBM graph-reading experts were also included in the Wagner et al. study). Participants completed think-alouds on two standard CBM graphs. Think-alouds were coded for accuracy, completeness, and sequential coherence, and for the number of...
statements comparing data across instructional phases (i.e., data-to-data comparisons), comparing data to the goal line or to the goal (i.e., data-to-goal comparisons), and linking data to instruction (i.e., data-to-instruction links).

Results revealed that teachers’ think-alouds were accurate, but were less complete and coherent than the CBM graph-reading experts’ think-alouds; however, teachers’ think-alouds were nearly as complete and coherent as those of the education graph-reading experts, and more complete and coherent than those of the general graph-reading experts. With regard to the “data-to” comparisons/links, teachers made fewer data-to-data and data-to-goal comparisons and fewer data-to-instruction links than the CBM graph-reading experts, but nearly the same number of such comparisons/links as the general graph-reading and education graph-reading experts.

The results of van den Bosch et al. (2017) were positive in the sense that they suggested that difficulties with CBM graph comprehension were not unique to teachers. However, the authors noted that there was a fair amount of variability in teachers’ completeness and coherence scores, revealing that some teachers struggled to describe CBM graphs in a complete and coherent manner. Of greater concern was the fact that teachers made few or none data-to-data and data-to-goal comparisons and data-to-instruction links. These comparisons/links represented teachers’ ability to read between and beyond the data; that is, to interpret the data and link it to instruction. Such skills are the essence of data-based decision-making. Van den Bosch et al. (2017) suggested that teachers might need specific, directed instruction on how to interpret CBM data and link it to instruction.

**Purpose of the Study: Improving Teachers’ Comprehension of CBM Progress Graphs**

In the present study we examine the effects of instruction on teachers’ CBM graph comprehension. We compare three different instructional approaches, each differing in the extent to which it emphasizes reading the data, interpreting the data, and linking the data to instruction. Instruction was delivered via online videos. We hypothesized that CBM instruction would lead to improvements in teachers’ CBM graph comprehension, and more important, that interactive instruction and practice on how to interpret CBM data and link it to instruction would result in better graph comprehension than basic instruction.

We also examine the social validity, or teachers’ acceptance, of both the CBM instruction and of CBM. Social validity is important to consider because educators’ attitudes towards data-based decision-making have been found to relate to the effects of data-based decision-making on student progress (Keuning, van Geel, & Visscher, 2017). We did not have a specific hypothesis regarding social validity, but speculated that differences across the three instructional approaches in terms of length and the
amount of teacher response required (see Method section) might lead to differences in teachers’ attitudes towards the CBM instructional videos and/or CBM.

Two research questions were addressed in the study:
1. What are the effects of CBM instruction on teachers’ CBM graph comprehension? Do effects differ by instructional approach?
2. What is the social validity of each instructional approach?

The study employed a randomized-control, pretest-posttest design, with the CBM instructional approach as the independent variable. CBM graph comprehension and social validity were the dependent variables. Analyses were conducted separately for each dependent variable.

**METHOD**

**Participants**
Participants were Dutch elementary-school teachers \( N = 164; \) 146 female, 18 male; \( M_{\text{age}} = 37.87 \) years, \( SD = 11.97, \) range 21-67) from 66 different schools. To participate, teachers had to have taught in grades 3 to 6 in the five years preceding the study. In the Netherlands, general education teachers are typically responsible for the instruction of students with dyslexia or students who struggle with reading. General education teachers organize what would, in the context of RTI, be considered Tier 1 and Tier 2 instruction for these students. Sometimes the students receive extra instruction from a specialist, but often the specialist is someone outside of the school system. In a limited number of cases, if problems are complex, students are sent to a specialized school. In our sample, most teachers \( n = 158 \) were general education teachers. A small number \( n = 6 \) worked in specialized schools. All teachers held a bachelor’s degree in education, and four also held master’s degrees in education, psychology, or languages. Teachers had an average of 13.44 years \( SD = 10.55, \) range 0-42) of teaching experience.

Teachers were recruited via convenience sampling, and were randomly assigned to either a control condition (CONTROL; \( n = 44 \)) or to one of three CBM instructional conditions: BASIC \( n = 38 \), INTERPRETATION \( n = 42 \), or INTERPRETATION + LINKING \( n = 40 \). Group size differences were due to the fact that 16 teachers were dropped from the original sample of 180 teachers (see following section). Demographic data for the four groups of teachers are provided in Table 5.1.
Table 5.1. Demographic Data for Teachers in the CONTROL, BASIC, INTERPRETATION, and INTERPRETATION + LINKING Conditions

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 44)</th>
<th>BASIC (n = 38)</th>
<th>INTERPRETATION (n = 42)</th>
<th>INTERPRETATION + LINKING (n = 40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (numbers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>41 (93.2%)</td>
<td>34 (89.5%)</td>
<td>35 (83.3%)</td>
<td>36 (90%)</td>
</tr>
<tr>
<td>Male</td>
<td>3 (6.8%)</td>
<td>4 (10.5%)</td>
<td>7 (16.7%)</td>
<td>4 (10%)</td>
</tr>
<tr>
<td>Age (in years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>38.66 (12.63)</td>
<td>38.47 (13.35)</td>
<td>36.88 (10.93)</td>
<td>37.48 (11.22)</td>
</tr>
<tr>
<td>Range</td>
<td>22-67</td>
<td>22-62</td>
<td>21-56</td>
<td>23-61</td>
</tr>
<tr>
<td>Type of school (numbers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General education</td>
<td>43 (97.7%)</td>
<td>36 (94.7%)</td>
<td>39 (92.9%)</td>
<td>40 (100%)</td>
</tr>
<tr>
<td>Special education</td>
<td>1 (2.3%)</td>
<td>2 (5.3%)</td>
<td>3 (7.1%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Highest degree (numbers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor's degree</td>
<td>44 (100%)</td>
<td>37 (97.4%)</td>
<td>40 (95.2%)</td>
<td>39 (97.5%)</td>
</tr>
<tr>
<td>Master's degree</td>
<td>0 (0%)</td>
<td>1 (2.6%)</td>
<td>2 (4.8%)</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td>Teaching experience (in years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>14.28 (11.52)</td>
<td>14.18 (11.93)</td>
<td>12.26 (9.03)</td>
<td>13.05 (9.76)</td>
</tr>
<tr>
<td>Range</td>
<td>1-42</td>
<td>1-42</td>
<td>1-36</td>
<td>0-36</td>
</tr>
</tbody>
</table>

Note: Demographic data for the 16 teachers that were dropped from the sample were: Gender: 13 females (81.3%), 3 males (18.8%); Age: M = 24.75, SD = 5.86, range 22-45; school type: 16 general education (100%); highest degree: 12 bachelor's (75%), 4 master's (25%); teaching experience: M = 2.66, SD = 4.28, range 0-17.
Teachers’ progress-monitoring experience

To provide a thorough description of the sample, and to examine group comparability on relevant factors, we collected data on teachers’ progress-monitoring experience and their general graph-reading skills. With regard to progress-monitoring experience, teachers were asked to answer 10 yes/no or open-ended questions asking about their experience using various progress-monitoring systems, including CBM. All elementary schools in the Netherlands are required to monitor students’ academic progress using nationally-normed standardized tests that are administered once or twice a year, and most schools use one of three commercial progress-monitoring systems to do so. Data from the systems are provided to teachers at both the individual and class level in the form of graphs and tables. CBM is not in widespread use in the Netherlands.

All teachers reported that their schools implemented one of the three Dutch progress-monitoring systems. Teachers were asked if they used the graphs from the systems and if so, what they did with the data from the graphs. Most teachers (n = 156 of 164) reported using the graphs, and reported using the data from the graphs to examine student progress, to communicate student progress to parents or colleagues, to place students in instructional groups, or to decide which students needed additional instruction.

Sixteen of the 180 teachers in the original sample reported that they had learned about CBM via University coursework, and that, as part of this coursework, they had interpreted data from CBM graphs and/or used CBM to monitor student progress. Despite random assignment, these 16 teachers were not evenly distributed across conditions (n = 2, 5, 4, and 5 in the CONTROL, BASIC, INTERPRETATION, and INTERPRETATION + LINKING conditions respectively). Demographic data for these 16 teachers (see Table 5.1, note) revealed that these teachers were younger and more highly educated, and had less teaching experience, than teachers in the final sample. Inspection of pretest scores revealed that pretest scores for the 16 teachers on the various aspects of CBM graph comprehension were in general higher than those for other teachers in the sample, especially with respect to scores related to interpreting and linking the data (see Results section). Based on this information, the decision was made to drop these 16 teachers from all further analyses. Of the remaining 164 teachers, none had ever interpreted CBM data or used CBM to monitor student progress.

Teachers’ general graph-reading skills

To examine general graph-reading skills, teachers completed two pretest-only measures: a Graph-reading Self-perception Scale and a Graph-reading Test. The Graph-reading Self-perception Scale was a translated version of the Graph-reading Ability subscale developed by Xi (2005). The scale included 12 items asking participants to rate their graph-reading
ability on a 6-point Likert scale, with 1 as low and 6 as high. Mean scale scores were nearly identical across the four groups (CONTROL: $M = 4.71$, $SD = 0.84$; BASIC: $M = 4.66$, $SD = 0.75$; INTERPRETATION: $M = 4.71$, $SD = 0.61$; INTERPRETATION + LINKING: $M = 4.74$, $SD = 0.90$). A Kruskall-Wallis Test revealed no significant between-group differences on self-perception scores, $\chi^2 (3) = 0.95, p = .81$.

The *Graph-reading Test* included 7 multiple-choice items requiring teachers to read and interpret line graphs. Items were based on items from a graph skills test developed by Shah and Freedman (2009). Teachers received one point per item answered correctly. Mean test scores were comparable across groups (CONTROL: $M = 3.64$, $SD = 1.06$; BASIC: $M = 3.32$, $SD = 1.12$; INTERPRETATION: $M = 3.74$, $SD = 1.11$; INTERPRETATION + LINKING: $M = 3.63$, $SD = 1.03$). A one-way ANOVA revealed no significant between-group differences on test scores, $F(3, 160) = 1.12, p = .34$.

**Independent Variable: CBM Instructional Approaches**

CBM instruction was delivered via instructional videos developed by the research team. Videos were used to ensure comparability in instruction across conditions. To create the videos, scripts for each of the three versions were written and audio-taped. Graphs, visualizations, and animations were created using Excel and PowToon. The spoken text, animations, and visualizations were combined using Adobe Première Pro. In the final step, interactive practice tasks were added to the INTERPRETATION and INTERPRETATION + LINKING videos.

The information in each video was presented by a narrator and illustrated via the story of a teacher, Mr. Kees, and his student, Sander, who had reading difficulties. The content of the videos was based on content from a university course on CBM, on a book written for practitioners on how to implement CBM (*The ABCs of CBM*; Hosp, Hosp, & Howell, 2007), and on training materials retrieved via the National Center on Progress Monitoring (www.studentprogress.org) and the Research Institute on Progress Monitoring (www.progressmonitoring.org). Each video began with an introduction that provided background on CBM. The introduction was followed by four segments on CBM implementation: collecting data, graphing data, interpreting data, and linking data to instruction. Differences between the three videos were seen in the segments on interpreting data and linking data to instruction. Differences are described in the following sections.

**BASIC condition**

In the BASIC condition, the segments on interpreting data and linking data to instruction consisted of explaining to and modeling for the teachers how to interpret the data and link it to instruction. The CBM instruction in this condition was not interactive and
Chapter 5

teachers did not practice the skills. Thus, in the segment on *interpreting data*, teachers were shown how an online progress-monitoring system provided recommendations to change instruction or raise the goal, and were told that the recommendations were based on the answers to three data-interpretations questions: (1) *Is the student making progress?* (2) *Will the student reach his or her goal?* and (3) *Does the instruction need to be changed?* Teachers were then shown six sample CBM graphs, each with one phase of data and a slope line drawn through the data. For each graph, they were shown the computer-generated instructional recommendation and were given an explanation of how it was generated.

In the segment on *linking data to instruction*, teachers were provided a description of five categories in which instructional changes could be made (i.e., activity, time, setting, material, and motivation), and were given an example of a potential change within each category (i.e., devoting more attention to a specific skill for *activity*, providing longer or more instruction for *time*, providing 1:1 instruction for *setting*, using materials from a different level for *material*, and using material that is of interest to the student for *motivation*). The BASIC video lasted 25 min.

**INTERPRETATION condition**

In the INTERPRETATION condition, teachers were given the same instruction as in the BASIC condition, but also were given interactive instruction and practice on interpreting the data. Thus, in the segment on *interpreting data*, for each of the six sample graphs, teachers themselves answered the three data-interpretation questions, and provided explanations for their answers. Teachers thus generated their own instructional decision before being shown (and explained) the recommendation that would be given by an online progress-monitoring system. Note that teachers did not receive specific, tailor-made feedback on the instructional decision that they made; they were, however, given the opportunity to compare their decision to the recommendation (and rationale) of the online system. This approach is similar to that used by L. S. Fuchs et al. (1989a). The INTERPRETATION video lasted approximately 35 min, depending on how long it took the teachers to answer the data-interpretation questions.

**INTERPRETATION + LINKING condition**

In the INTERPRETATION + LINKING condition, teachers were given the same instruction and interactive practice on interpretation as was given to teachers in the INTERPRETATION condition. However, in the segment on *linking data to instruction*, teachers were given additional interactive practice. Specifically, teachers were asked to provide “advice” to the teacher in the video (Mr. Kees) about selecting an instructional change. Following the description of the five categories of instructional changes given
in the BASIC and INTERPRETATION conditions, Mr. Kees reflected upon five possible changes he could make (one for each category). Teachers were then asked to select one of the five changes for Mr. Kees to implement, and were asked to provide a rationale for their selection. They selected changes two times. For this practice, no feedback was provided to teachers. The INTERPRETATION + LINKING video lasted approximately 45 min, depending on how long it took the teachers to answer the data-interpretation questions and to select instructional changes.

**CONTROL condition**

In the CONTROL condition, instead of watching the CBM instructional video, teachers were provided with two “filler” tasks, including a short video clip about assessment in schools and a survey asking their opinions about this topic. After study completion, for ethical reasons, CONTROL teachers were shown the BASIC instructional video.

**Primary Dependent Variable: Teachers’ CBM Graph Comprehension**

The primary dependent variable in the study was teachers’ CBM graph comprehension. At both pre- and posttest, graph comprehension was assessed via a CBM Graph-description Task in which teachers examined a CBM graph for 1 min, and then described the graph as if they were describing it to the student’s parent. No time limits were placed on the graph descriptions. Note that the prompt used in the present study was different than that used in previous CBM graph-comprehension studies. In previous studies, teachers were asked to tell all they were seeing and thinking (Espin et al., 2017; van den Bosch et al., 2017; Wagner et al., 2017). As suggested in van den Bosch et al., we used a more “realistic” prompt for the task.

The graphs used for the Graph-description Task were two researcher-made CBM graphs that depicted progress of a fictitious student in reading across one school year (see Figure 5.1, for a sample graph). Data points on the graph reflected the student’s scores on maze-selection tasks. The two graphs were parallel in format and graph patterns. For example, although the actual data points differed across the two graphs, in both graphs, the slope line for Phase 1 began above, and converged upon, the goal line. The patterns displayed in the graph were CBM graph patterns that had been found in previous research to be difficult to interpret (Espin, Saab, Pat-El, Boender, & van der Veen, 2018). One graph was administered at pretest and the other at posttest. The order was counterbalanced across teachers. Prior to completing the pretest, teachers were shown a sample CBM graph, and were provided a general description of the graph (see van den Bosch et al., 2017, for description). In addition, they were shown a maze probe, and were told how the maze task was used within CBM.
CBM Graph-description Coding

Graph-description coding procedures were adapted from those used in previous research (Espin et al., 2017; van den Bosch et al., 2017; Wagner et al., 2017). Teachers' graph descriptions were audiotaped and transcribed, and each transcription was checked by a second coder. Graph descriptions were then parsed into idea units (i.e., statements that expressed one idea), and each idea unit was assigned a content code corresponding to one of eight graph elements: Framing (statements describing the set-up of the graph such as titles, axes, and the legend), Baseline (statements describing baseline data of the student/peers or procedures to obtain those data), Goal Setting (statements describing the goal or procedures to set the goal), Instructional Phases 1, 2, 3, and 4 (statements describing the data within a particular phase), and Goal Achievement (statements describing whether the student achieves the goal), or were assigned the code General Progress (statements describing the data across the four phases) or Reflective Statement (statements reflecting upon or evaluating the graph content).

Following the assignment of content codes, the graph descriptions were coded in two separate rounds of coding. In the first round, the descriptions were coded for accuracy, completeness, sequential coherence, and specificity. Accuracy was the percentage of statements correctly reflecting the data presented on the graph. Completeness was the number of graph elements mentioned (of the eight elements listed earlier). Sequential coherence was the percentage of statements that followed a logical and coherent sequence. To calculate sequential coherence, teachers' descriptions were compared to an “ideal” sequence, that is, a sequence in which the eight graph elements were described in an order that reflected CBM implementation, and that progressed from Framing to Baseline to Goal Setting to Instructional Phases 1, 2, 3, and 4 to Goal Achievement (see Espin et al., 2017, for a detailed description of the coding process).

Specificity was the percentage of statements that referred to progress in a specific instructional phase (coded as a Phase 1, 2, 3, or 4 statements) as opposed to statements that described progress in general (coded as a General Progress statement). Specificity was calculated by dividing the number of statements made about specific instructional phases by the total number of progress statements made. Note that the denominator we used for specificity is different than that used in Espin et al. (2017) and Wagner et al. (2017). In these studies, the denominator was the total number of statements included in the graph description.

In a second round of coding, data-to-data comparisons, data-to-goal comparisons, data-to-instruction links, and raising-the-goal comments were coded. Data-to-data comparisons were the number of times teachers compared student performance or progress data in one instructional phase to either the baseline phase or to another instructional phase (e.g., His scores in Phase 1 are higher than his baseline scores. or He shows more progress in
Phase 2 than in Phase 1). Data-to-goal comparisons were the number of times teachers compared student performance or progress data to the goal line or to the goal (e.g., Her scores are all below the goal line or At this rate of growth she will achieve her goal). Data-to-instruction links were the number of times teachers linked student performance or progress data to the student’s reading instruction and included comments that referred to the effectiveness of instruction or to the need for instructional changes (e.g., He is making progress in Phase 3, thus the change in instruction was effective or His slope line was flat so the instruction was changed). Finally, raising-the-goal comments were the number of times teachers stated that the student’s goal should be raised when progress was greater than expected (e.g., Her slope is steeper than the goal, so the goal should be raised).

Graph descriptions were coded by the first author and 10 master’s students in Education and Child Studies. The students were trained by the first author across a number of sessions, each focusing on the various aspects of the coding procedures. Sessions lasted from 10-60 min, depending on which aspect was being addressed in the session. At the end of each session, the master’s students coded sample descriptions, and agreement with the first author was calculated. Students had to reach 80% agreement before they could begin coding that aspect. All data were double coded by the first author and one master’s student. Coding disagreements were discussed and resolved. If agreement could not be reached the second author was consulted.

Intercoder agreement
Intercoder agreement was calculated for every third graph description prior to discussions between coders. For the content codes (Framing, Baseline, Goal Setting, etc.) and for accuracy, agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements, and multiplying by 100. Agreement for the content codes was 89.94%. (Recall that content codes were used to calculate completeness, sequential coherence, and specificity; therefore, separate agreement percentages are not reported for these variables). Agreement for accuracy was 97.03%.

For data-to-data comparisons, data-to-goal comparisons, data-to-instruction links, and raising-the-goal comments, coders identified the number of occurrences of each variable in the graph descriptions. Agreement was calculated per variable in two steps. First, agreement was calculated per graph description by dividing the smaller number of occurrences by the larger number of occurrences, and multiplying by 100. Second, average agreement across descriptions was calculated. Agreement was 81.63% for data-to-data comparisons, 87.46%, for data-to-goal comparisons, 80.90% for data-to-instruction links, and 98.36% for raising-the-goal comments.
Secondary Dependent Variable: Social Validity of CBM Instructional Approaches

The social validity of each CBM instructional approach was assessed via a self-developed 10-item Social Validity Scale. Teachers first rated five statements about CBM (*I understand what CBM is; I think CBM graphs are easy to read; I think CBM graphs would be helpful for instructional decision-making; I think I am sufficiently trained to use CBM in my class; I would like to use CBM in my class for individual students with reading problems.*), and four statements about the instructional video (*I thought the CBM video instruction was clear/interesting/useful/informative*). These 9 items were on a 4-point Likert scale, ranging from *strongly disagree* to *strongly agree*. The tenth item asked teachers to rate the video on a scale of 1-10, with 10 being the highest rating. (As described earlier, CONTROL teachers were shown the BASIC instructional video after study completion. They, then, also completed the Social Validity Scale for this video, thus Social Validity data were also available from the CONTROL group.)

Procedures

Data for the study were collected on an individual basis at a place convenient for the teacher (school, home, or university) in a session lasting from 1.5 to 2 hrs. Teachers in the three CBM instructional conditions completed the tasks in the same order: Graph-reading Self-perception Scale, Graph-reading Test, Pretest CBM Graph-description Task, CBM instructional video, Social Validity Scale, Posttest CBM Graph-description Task, and demographic questionnaire. Teachers in the CONTROL condition completed the tasks in nearly the same order, with the exception that, instead of watching the CBM instructional video, teachers completed the two filler tasks about assessment in schools. In addition, teachers completed the demographic questionnaire at this point. After study completion control teachers were shown the BASIC instructional video and completed the Social Validity Scale for this video.

Data collectors were the first author and the 10 master’s students. Prior to data collection, the students were trained by the first author in a single session in which they practiced all data collection procedures. With the exception of the Pre- and Posttest CBM Graph-description Task, all data were collected via computer. The CBM Graph-description Task was administered by the data collector, who gave instructions and acted as the “parent” who listened to the graph description. Data collectors were present during the entire session to ensure that the teachers completed all tasks, and to collect fidelity data. All teachers completed all tasks except for one teacher, who did not complete the Social Validity Scale. Fidelity of implementation for the CBM instructional videos was 100%; that is, all participating teachers watched the correct video and all teachers completed all interactive practice tasks during the video. Data collectors were observed by the first author on their first data-collection session, and again at (approximately) the 15th session. All data collectors adhered to all data-collection procedures.
RESULTS

Teachers’ CBM Graph Comprehension: Descriptives

Descriptive statistics on teachers’ pre- and posttest scores on the various aspects of CBM graph comprehension are reported in Table 5.2. Pretest scores for the 16 teachers that were dropped from the original sample are reported in the note under the table. It is worthwhile to note that including these 16 teachers in the analyses did not change the pattern of results that were found.

Table 5.2. Means and SD’s (in parentheses) from the Pre- and Posttest CBM Graph-description Task

<table>
<thead>
<tr>
<th></th>
<th>Total sample (N=164)</th>
<th>CONTROL (n=44)</th>
<th>BASIC (n=38)</th>
<th>INTERPRETATION (n=42)</th>
<th>INTERPRETATION + LINKING (n=40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre ACC (%)</td>
<td>97.87 (4.44)</td>
<td>97.70 (5.14)</td>
<td>96.52 (7.31)</td>
<td>97.66 (3.59)</td>
<td>98.90 (2.58)</td>
</tr>
<tr>
<td>Post ACC (%)</td>
<td>98.81 (5.98)</td>
<td>97.70 (4.09)</td>
<td>96.52 (7.24)</td>
<td>95.40 (7.31)</td>
<td>97.78 (2.58)</td>
</tr>
<tr>
<td>Pre COM (#)</td>
<td>7.18 (1.30)</td>
<td>7.34 (1.50)</td>
<td>6.80 (1.59)</td>
<td>7.24 (1.42)</td>
<td>7.50 (0.83)</td>
</tr>
<tr>
<td>Post COM (#)</td>
<td>7.34 (1.15)</td>
<td>6.80 (1.59)</td>
<td>7.24 (1.42)</td>
<td>7.34 (1.12)</td>
<td>7.62 (0.83)</td>
</tr>
<tr>
<td>Pre SEQ (%)</td>
<td>55.20 (19.89)</td>
<td>64.75 (17.12)</td>
<td>59.24 (20.22)</td>
<td>55.79 (22.09)</td>
<td>65.95 (21.81)</td>
</tr>
<tr>
<td>Post SEQ (%)</td>
<td>64.75 (17.12)</td>
<td>59.24 (20.22)</td>
<td>55.79 (22.09)</td>
<td>65.95 (21.81)</td>
<td>60.58 (14.69)</td>
</tr>
<tr>
<td>Pre SPEC (%)</td>
<td>81.65 (28.97)</td>
<td>92.91 (20.17)</td>
<td>74.45 (33.72)</td>
<td>79.69 (32.14)</td>
<td>95.13 (31.41)</td>
</tr>
<tr>
<td>Post SPEC (%)</td>
<td>92.91 (20.17)</td>
<td>74.45 (33.72)</td>
<td>79.69 (32.14)</td>
<td>95.13 (31.41)</td>
<td>95.40 (22.09)</td>
</tr>
<tr>
<td>Pre DD (#)</td>
<td>1.35 (0.98)</td>
<td>1.30 (1.19)</td>
<td>1.34 (1.22)</td>
<td>1.36 (1.22)</td>
<td>1.36 (1.22)</td>
</tr>
<tr>
<td>Post DD (#)</td>
<td>1.30 (1.08)</td>
<td>1.34 (1.19)</td>
<td>1.36 (1.22)</td>
<td>1.36 (1.22)</td>
<td>1.39 (1.19)</td>
</tr>
<tr>
<td>Pre DG (#)</td>
<td>1.57 (1.19)</td>
<td>1.50 (1.08)</td>
<td>1.50 (1.14)</td>
<td>1.50 (1.14)</td>
<td>1.50 (1.14)</td>
</tr>
<tr>
<td>Post DG (#)</td>
<td>1.50 (1.08)</td>
<td>1.50 (1.08)</td>
<td>1.50 (1.14)</td>
<td>1.50 (1.14)</td>
<td>1.50 (1.08)</td>
</tr>
</tbody>
</table>

Note. ACC = Accuracy, COM = Completeness (out of 8), SEQ = Sequential coherence, SPEC = Specificity, DD = Data-to-data comparisons, DG = Data-to-goal comparisons, DI = Data to-instruction links. Mean pretest scores for the 16 teachers that were dropped from the sample were: ACC: 97.71 (SD = 4.55), COM: 7.63 (SD = 0.81), SEQ: 59.97 (SD = 14.95), SPEC: 95.78 (SD = 7.73), DD: 2.31 (SD = 1.25), DG: 2.63 (SD = 1.89), DI: 4.88 (SD = 2.50).

Mean pretest scores for the total sample (see column 1 in Table 5.2) revealed that teachers’ graph descriptions were accurate and complete prior to CBM instruction. Approximately half of the statements that teachers made in their descriptions were in a logical, coherent order, and a large percentage of their statements were specific. Teachers made data-to-data comparisons, data-to-goal comparisons, and data-to-instruction links. Mean posttest scores for the total sample revealed that scores increased for sequential coherence, specificity, data-to-goal comparisons, and data-to-instruction links.
links. Dependent t-tests and Wilcoxon signed-rank tests revealed that these increases were statistically significant (sequential coherence: $t(163) = -5.44, p < .001, d = -0.42$, specificity: $z = -5.31, p < .001, r = -.29$, data-to-goal comparisons: $z = -5.88, p < .001, r = -.32$, data-to-instruction links: $z = -8.52, p < .001, r = -.47$). No significant increases were found for accuracy ($z = -0.31, p = .75$), completeness ($z = -1.2, p = .23$), or data-to-data comparisons ($t(163) = -0.62, p = .54$).

Comments about raising the goal were rarely made in the think-alouds and thus means and standard deviations are not reported in Table 5.2. Instead it was counted how many teachers mentioned raising the goal. At pretest, no teacher mentioned raising the goal. At posttest, only 30 teachers (18%) mentioned raising the goal: 0 in the CONTROL, 6 (16%) in the BASIC, 11 (26%) in the INTERPRETATION, and 13 (33%) in the INTERPRETATION + LINKING groups. The significance of the differences in percentages between groups was examined using six pairwise “N-1” Chi-squared tests for comparisons of proportions with a Bonferroni corrected alpha level of .0083 (0.05/6) for each test. Results revealed significant differences between the three CBM instructional groups and the CONTROL group ($\chi^2(1) = 7.51, p = .0061$ for BASIC, $\chi^2(1) = 12.95, p = .0003$ for INTERPRETATION, and $\chi^2(1) = 17.02, p < .0001$ for INTERPRETATION + LINKING). No significant differences were found between the three CBM instructional groups.

**Effectiveness of CBM Instructional Approaches: Profile Analysis**

To examine pre-post group differences in teachers’ CBM graph comprehension, profile analysis was conducted. Profile analysis is a particular way of conducting multivariate analysis of variance (MANOVA) and often is described as the multivariate approach to repeated measures analyses (Tabachnick & Fidell, 2012). Profile analysis is used to compare the profiles of groups of participants who have been measured on several dependent variables at the same time. The results of the analysis are presented visually in the form of profile plots that represent the profiles of different groups of participants for a set of dependent variables.

When comparing the profiles of different groups of participants, three types of questions are answered: (1) *Are the profiles of the different groups parallel?*, (2) *Does one group, on average, have higher scores for the dependent variables than another group?*, and (3) *Are the scores for all of the dependent variables on average the same?* These three questions are referred to as *parallelism, difference in levels,* and *flatness,* and are answered by examining the GLM results of the within-subjects interaction effect, the between-subjects effect, and the within-subjects effect, respectively. If the profiles are not parallel, the question of flatness is irrelevant because non-parallel profiles are per definition not flat.
The goal of our profile analysis was to compare the CBM graph-comprehension profiles for teachers in the four conditions: CONTROL, BASIC, INTERPRETATION, and INTERPRETATION + LINKING. The aspects of comprehension included in the analysis were sequential coherence, specificity, data-to-data comparisons, data-to-goal comparisons, and data-to-instruction links. Accuracy and completeness were not included because there was a ceiling effect and relatively little variation in these scores at both pre- and posttest (see Table 5.2). Raising-the-goal comments were not included because of their relatively low occurrence.

To compare groups in a profile analysis, all variables must be on the same scale (Tabachnick & Fidell, 2012), thus, scores on our variables had to be standardized. We calculated pre-post change scores by subtracting the pretest from the posttest scores, and we then standardized these change scores.

The first step in profile analysis was to evaluate assumptions. The groups had sufficiently equal sample sizes, the smallest group \((n = 38)\) included (far) more cases than the number of dependent variables (5), justifying multivariate analysis, and there were no missing data. The assumption of multivariate normality was met for all dependent variables except specificity. The distribution of specificity was bimodal: For most teachers, their pre- and posttest specificity scores were relatively similar, but for a few teachers there was a large discrepancy between pre- and posttest scores (for example, at pretest, 0% of their statements were specific, whereas at posttest 100% of their statements were specific). Given that the sample size was sufficiently large \((n > 30)\), specificity was kept in the analysis.

Inspection of boxplots revealed 13 outliers for data-to-goal comparisons and 12 outliers for specificity. We had no valid reason to drop the outliers from the sample and for all cases, Cook’s distances were smaller than 1, indicating that the outliers had no substantial impact in determining the outcome of the analysis. Thus, the outliers were kept in the data set. The assumptions for linearity, homogeneity of regression, and multicollinearity and singularity were met. The assumption of homogeneity of variance-covariance matrices was not met; therefore Games-Howell corrections were used to interpret the results.

A GLM repeated measures analysis with teachers’ standardized pre-post change scores for CBM graph comprehension was conducted. Results of the within-subjects effect test for parallelism (i.e., the interaction between CBM graph comprehension and condition) revealed that the profiles of teachers in the four groups were significantly non-parallel, indicating that the profile of the dependent variables differed for at least one group (see Figure 5.2; \(F(12, 640) = 6.29, p > .001, \eta_p^2 = .11\) (medium effect), 95% CI = \(-1.26, 0.69\).
Results of the between-subject effect to test for the differences in levels revealed a main between-subjects effect for condition ($F(3, 160) = 6.85, p < .001, \eta^2_p = .11$ (medium effect), 95% CI = -0.44, 0.31). Multiple follow-up comparisons with Games-Howell correction revealed that teachers in the three CBM instructional groups had larger pre-post change scores for the total set of dependent variables than teachers in the CONTROL group ($p = .007$ for BASIC, $p = .001$ for INTERPRETATION, and $p = .004$ for INTERPRETATION + LINKING). No significant differences in pre-post change scores were found among the three instructional groups. Further, the mean pre-post change score for the total set of dependent variables represented an increase for teachers in the instructional groups, whereas it represented a decrease for the CONTROL teachers (see Figure 5.2). The flatness test was not examined because profiles were not parallel.

The deviation from parallelism was evaluated by examining the profiles per group (i.e., the group profiles) through simple-effects analyses. That is, we examined group differences for each dependent variable. The simple effects analyses revealed that there were no differences between groups in pre-post change scores for sequential coherence ($F(3, 160) = 1.07, p = .37$), specificity ($F(3, 160) = 0.49, p = .69$), or data-to-data comparisons ($F(3, 160) = 1.17, p = .32$).
However, differences between groups were found for data-to-goal comparisons ($F(3, 160) = 6.14, p = .001$, $\eta^2_p = .10$ (medium effect)) and for data-to-instruction links ($F(3, 160) = 33.49, p < .001$, $\eta^2_p = .39$ (large effect)). Contrasts were defined based on an evaluation of the profile plot (see Figure 5.2), where it can be seen that for data-to-goal comparisons (DG), teachers in the three instructional groups had larger pre-post changes scores than CONTROL teachers (contrast 1). Moreover, it can be seen that teachers in the INTERPRETATION group had larger pre-post change scores than teachers in the BASIC group (contrast 2), and that teachers in the INTERPRETATION + LINKING group had larger pre-post changes scores than teachers in the INTERPRETATION group (contrast 4). The first three contrasts were significant ($t(160) = -2.61, p = .01, d = -0.41$; $t(160) = 2.31, p = .02, d = 0.37$; and $t(160) = 3.25, p = .001, d = 0.51$ respectively), but the fourth contrast was not ($t(160) = 0.84, p = .40$). With regard to data-to-instruction links (DI), the profile plot revealed that teachers in the three instructional groups had larger pre-post changes scores than CONTROL teachers (see Figure 5.2), and this contrast was significant ($t(160) = -3.86, p < .001, d = -0.61$).

In sum, the results of the profile analysis revealed that the profiles of the CBM graph-comprehension pre-post change scores were not parallel across the four groups. Overall, teachers in the three CBM instructional groups had significantly larger pre-post change scores for CBM graph comprehension than did teachers in the CONTROL group. Further, the mean pre-post change score for teachers in each of the three instructional groups across the total set of dependent variables was positive, whereas for teachers in the CONTROL group it was negative. More specifically, teachers who received CBM instruction improved more on data-to-goal comparisons and data-to-instruction links than teachers who received no CBM instruction. Moreover, for data-to-goal comparisons, significant differences were found among the instructional groups, with the teachers in the INTERPRETATION and the INTERPRETATION + LINKING groups improving more than teachers in the BASIC group.

Social Validity of CBM Instructional Approaches: Descriptives.

Teachers’ overall ratings for the CBM instructional videos were fairly positive, with an average rating across teachers of 7.96 ($SD = 0.96$) out of 10. Mean overall ratings were similar across the four groups (BASIC: $M = 7.82$ ($SD = 1.18$); INTERPRETATION: $M = 7.79$ ($SD = 0.90$); INTERPRETATION + LINKING: $M = 8.05$ ($SD = 0.79$); CONTROL: $M = 8.18$ ($SD = 0.76$)). (Recall that CONTROL teachers rated the BASIC video after study completion.) Teachers’ ratings of the specific characteristics of the videos (i.e., clear, interesting, useful, informative) also were positive, with an average rating across teachers of 3.52 ($SD = 0.46$) out of 4. Mean ratings were similar across the four groups (BASIC: $M = 3.43$ ($SD = 0.48$); INTERPRETATION: $M = 3.57$ ($SD = 0.45$); INTERPRETATION + LINKING: $M = 3.53$.
Ratings of CBM also were positive, with an average rating across teachers of 3.54 ($SD = 0.38$) out of 4. Mean ratings were similar across the four groups (BASIC: $M = 3.48$ ($SD = 0.38$); INTERPRETATION: $M = 3.47$ ($SD = 0.40$); INTERPRETATION + LINKING: $M = 3.58$ ($SD = 0.35$); CONTROL: $M = 3.61$ ($SD = 0.37$)).

**DISCUSSION**

Reading and interpreting CBM progress graphs is an important skill for general and special education teachers. CBM progress graphs often provide focus for team meetings, and are meant to guide instructional decision-making for students in Tier 2 and Tier 3. Yet, if teachers do not interpret the data correctly and/or do not link it to instruction, the data are useless. The purpose of this study was to examine the effects of instruction on teachers’ CBM graph comprehension, and to examine whether effects differed by instructional approach. We also examined the social validity of each instructional approach.

**Effectiveness of CBM Instructional Approaches**

As a group, teachers became more coherent and specific in their graph descriptions, and their descriptions included more data-to-goal comparisons and data-to-instruction links at posttest than at pretest. Improvements in CBM graph comprehension from pre- to posttest were significantly greater for teachers in the three CBM instructional groups than for teachers in the control group, and these differences were due primarily to the improvements in data-to-goal comparisons and data-to-instruction links. That is, teachers in the three instructional groups were more likely to compare student performance or progress to the goal (line) and to describe the link between that performance or progress and instruction, than were teachers in the control group.

Improvements in data-to-goal comparisons and data-to-instruction links are of special importance within CBM. As mentioned earlier, it are these aspects of graph comprehension that have proven to be challenging for teachers (van den Bosch et al., 2017) but are the very essence of CBM data-based decision-making. To make CBM data-based decisions, teachers must compare student performance or progress to the goal to determine whether the progress is as expected/desired (data-to-goal comparison), and then must link the information to instruction to decide whether there is a need to change instruction (data-to-instruction link). The improvements that we found for data-to-instruction links especially are impressive when one considers that the number of such links on the posttest of teachers in the instructional groups was 4.9, which was
nearly the same as the number made by the CBM graph-reading experts in van den Bosch et al. (2017), who made 5 such links.

Although it is interesting that teachers in the instructional groups improved more than teachers in the control group on certain aspects of CBM graph comprehension, our primary interest was whether differences among the three instructional approaches would be found. Results of the profile analysis revealed that teachers in both the INTERPRETATION and INTERPRETATION + LINKING groups showed a greater increase in data-to-goal comparisons than teachers in the BASIC group, suggesting that the interactive instruction and practice included in the INTERPRETATION and INTERPRETATION + LINKING videos was worthwhile, at least with regard to making data-to-goal comparisons.

It was somewhat surprising that no significant differences were found for the other aspects of CBM graph comprehension examined in this study. With regard to accuracy and completeness, scores were high at pretest for all groups, leaving little room for improvement. On average, teachers were nearly 98% accurate, and were complete in their graph descriptions, mentioning 7 of 8 possible graph elements (88%). The completeness scores are higher than those found in van den Bosch et al. (2017), where teachers mentioned only 6 of 9 graph elements (67%). Differences may relate to the fact that we asked teachers to describe the graphs as if they were describing them to a parent, whereas van den Bosch et al. asked teachers to tell all they were seeing and thinking. Perhaps the parent-directed prompt led teachers to be more complete in their descriptions. With regard to data-to-data comparisons, it was not clear why teachers did not improve. That is, there was room for improvement. It could be that comparing student data from one instructional phase to data from another phase is a difficult skill and/or that teachers are not used to comparing data across phases, which is quite unique to CBM. Teachers, therefore, may need more or more explicit instruction on how to make data-to-data comparisons and on the importance of making such comparisons than was provided in the CBM instructional videos.

With regard to sequential coherence and specificity, results revealed that practice alone was enough to improve teachers’ coherence and specificity in graph descriptions. However, it is important to note that for sequential coherence there was room for more improvement. Sequential coherence for teachers on the posttest was only 65%. For the CBM experts in the Wagner et al. (2017)/van den Bosch et al. (2017) studies, sequential coherence was 85%. It was disappointing that teachers in the instructional groups did not improve more in describing CBM graphs in a logical and sequential manner than control teachers. Such a skill is important for communicating with parents in parent meetings or with professionals in team meetings. Teachers may need more than a relatively short (25-45 min) instructional video to describe CBM graphs in coherent
manner, or may need explicit instruction on how to effectively communicate CBM data. Such explicit instruction was not included in the CBM instructional videos.

Although because of their low occurrence, raising-the-goal comments were not included in the profile analysis, we think it is important to reflect upon the percentage of teachers who made such comments. At pretest, no teacher mentioned raising the goal. At posttest, teachers in all three instructional groups, but not in the control group, mentioned raising the goal, with the largest percentage seen in the INTERPRETATION and INTERPRETATION + LINKING groups. L. S. Fuchs, Fuchs, and Hamlett (1989b) found that teachers who set more ambitious goals and who raised the goal in response to the CBM data affected greater student gains than teachers who set less ambitious goals and did not raise the goal in response to the data. Our results suggest that CBM video instruction, and in particular the two videos that provide interactive instruction and practice, may serve to raise teachers’ awareness of the importance of raising the goal. However, it will be imperative to examine whether effects are seen in actual CBM implementation, and it should be noted that even for the most extensive instructional approach still two-thirds of the teachers did not mention raising the goal at posttest.

Social Validity of CBM Instructional Approaches
Teachers’ positive evaluations of the CBM instructional videos and of CBM itself supported the social validity of all three CBM instructional approaches. Given that educators’ positive attitudes towards data-based decision-making are related to the effects of data-based decision-making on student progress (Keuning et al., 2017), it is encouraging that the teachers, who were not familiar with CBM prior to participating in the study, developed a positive attitude about CBM via the instructional videos, and were positive about the video instruction itself.

Our results related to social validity fit well with results from Kennedy et al. (2016), who found that preservice teachers who received multimedia CBM instruction were more motivated than those who received the same instruction in an article format. Such results provide tentative support for the use of technology to provide CBM instruction to teachers. Such instruction can easily be incorporated into an online progress-monitoring system, or offered as a stand-alone course in the context of e-learning.

Limitations and Related Directions for Future Research
A limitation of the study was that the differences between the three instructional approaches may not have been large enough. For example, even in the BASIC condition, teachers received fairly detailed instruction on how to interpret and link the CBM data to instruction. It is not clear to what extent “typical” CBM instruction provides such instruction. It would be worthwhile to conduct a study to examine the content and form...
of “typical” CBM instruction.

A second limitation of the study was that it included a convenience sample of teachers. Replication of the study with other samples of teachers is in order.

Conclusion and Implications
In conclusion, our results suggest that CBM video instruction can be used to improve two important aspects of teachers’ CBM graph comprehension: Making data-to-goal comparisons and linking data to instruction. Given that both the INTERPRETATION and INTERPRETATION + LINKING approaches resulted in significant improvements in one of these areas (making data-to-goal comparisons) over the BASIC condition, we could suggest use of the INTERPRETATION approach to instruction, which is shorter than an INTERPRETATION + LINKING approach. However, we present our conclusions with caution because we did not examine the effects of each instructional approach on teachers’ actual CBM implementation or data-based decision-making, and on the performance of students who struggle with reading. Before firm conclusions are drawn about the effects of various approaches to CBM instruction, it will be essential to examine the effects of each approach on actual CBM implementation and data-based decision making, and on student performance.