

Data-Based Decision-Making: Teachers' Comprehension of Curriculum-Based Measurement Progress-Monitoring Graphs

Roxette M. van den Bosch, Christine A. Espin, and Siuman Chung

Leiden University

Nadira Saab

Leiden University Graduate School of Teaching (ICLON)

Teachers have difficulty using data from Curriculum-based Measurement (CBM) progress graphs of students with learning difficulties for instructional decision-making. As a first step in unraveling those difficulties, we studied teachers' comprehension of CBM graphs. Using think-aloud methodology, we examined 23 teachers' ability to *read*, *interpret*, and *link CBM data to instruction* for fictitious graphs and their own students' graphs. Additionally, we examined whether graph literacy—measured with a self-report question and graph-reading skills test—affected graph comprehension. To provide a framework for understanding teachers' graph comprehension, we also collected data from “gold-standard” experts. Results revealed that teachers were reasonably proficient at *reading the data*, but had more difficulty with *interpreting* and *linking the data to instruction*. Graph literacy was related to some but not all aspects of teachers' CBM graph-comprehension ability. Implications for training teachers to comprehend and use CBM progress data for decision-making are discussed.

Teachers are problem solvers. They are confronted each day with solving the problem of how best to help children learn. Teachers of students with learning difficulties face special challenges in their problem solving efforts. First, students with learning difficulties may not respond to the type of instructional approaches found to be effective for other students. Second, students with learning difficulties may improve at slow, incremental rates, yet instructional time is limited. Teachers cannot afford to waste precious educational time on interventions that are ineffective. To be successful problem solvers, teachers of students with learning difficulties must be relentless in their instruction. They must teach their students with a sense of urgency, striving to build increasingly effective instructional programs (Zigmond, 1997, 2003).

One important tool for building effective instructional programs for students with learning difficulties is a database of effective instructional interventions (e.g., *What Works Clearinghouse*, see <http://ies.ed.gov/ncee/wwc/>). Yet, students respond differentially to interventions—even to those with an empirical evidence base (Deno, 1985; Deno & Fuchs, 1987). Therefore, teachers must have a second tool available, one that allows them to collect data on the effectiveness of interventions for individual students. Furthermore, teachers must have the skills needed to use the data generated by such a tool to inform their instruction. One such assessment tool that teachers can use to evaluate the effects of instructional

programs on student progress is Curriculum-based Measurement (CBM; Deno, 1985).

Curriculum-Based Measurement

CBM is a progress-monitoring system designed to track the progress of individual students with learning difficulties, and to evaluate the effectiveness of instruction for those students (Deno, 1985, 2003). CBM involves frequent (e.g., weekly) administration of short, simple measures that sample performance in an academic area such as reading. Scores from the measures are placed on a graph that depicts student performance and progress over time. Key components of a CBM progress-monitoring graph include: (1) baseline data, representing the student's current level of performance; (2) peer data, representing typical performance and reflecting the discrepancy between the student and peers; (3) a goal line, representing the expected rate of growth and end-of-year level of performance; (4) data points, representing the number of correct and incorrect responses on weekly probes; (5) slope or growth lines, representing the student's rate of growth over time; and (6) solid vertical lines, representing instructional changes (see Figure 1 for a sample CBM graph).

In order to evaluate the effectiveness of instruction for a particular student, the teacher examines the graph to determine whether the student is progressing at the desired rate of growth and whether the student will achieve the goal. If growth is greater than expected, the teacher raises the goal. If growth is less than expected, the teacher changes instruction and then continues to monitor to examine the effects

Requests for reprints should be sent to Roxette van den Bosch, Leiden University. Electronic inquiries should be sent to r.m.van.den.bosch@fsw.leidenuniv.nl

Curriculum-Based Measurement

Maze scores of a 5th grade student with reading difficulties

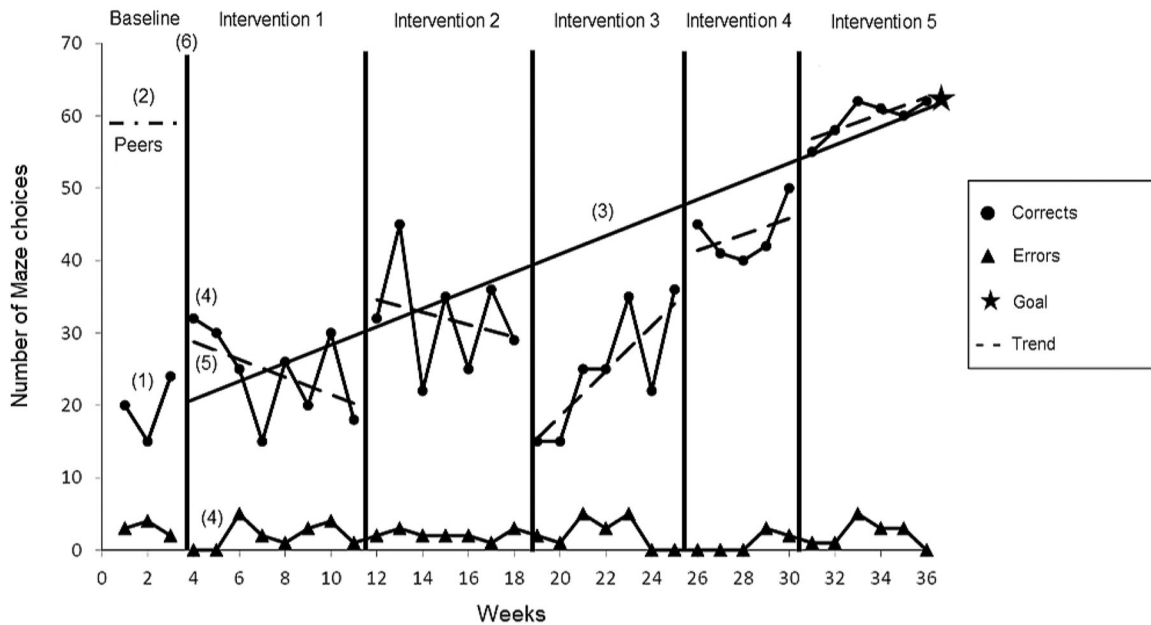


FIGURE 1 Sample of a standard CBM graph. Graphs were presented to participants in Dutch. Numbers were added to this sample graph for illustrative purposes: (1) = baseline data, (2) = peer data, (3) = goal line, (4) = data points, (5) = slope or growth line, and (6) = solid vertical line.

of the change. By responding to student data with goal or instructional changes, the teacher strives to build a powerful, effective instructional program for the student.

A large body of research has addressed the development of CBM measures and procedures in areas such as math, writing, and reading (see Foegen, Jiban, & Deno, 2007; McMaster & Espin, 2007; Wayman, Wallace, Wiley, Tichá, & Espin, 2007, for reviews), and demonstrates that *when* teachers use CBM to inform their instruction, student achievement improves (Stecker, Fuchs, & Fuchs, 2005). However, this research also reveals that teachers often do *not* use CBM to inform their instruction; that is, teachers collect and graph the data, but do not respond to the data with goal or instructional changes. To address this concern, Fuchs, Fuchs and colleagues developed approaches for, and investigated the effects of, computer-assisted feedback on CBM data-based decision-making (see Fuchs & Fuchs, 2002; Stecker et al., 2005, for reviews). They did not, however, examine teachers' understanding or interpretation of CBM progress graphs.

Graph Comprehension

The first step in CBM data-based decision-making is to interpret the progress graph—that is, to determine whether the graph signals the need for a goal or instructional change. At first glance, CBM graphs seem easy to interpret. After all, the graphs are designed to be simple, clear, and easy to understand (Deno, 1985, 2003); however, research suggests that graph interpretation is not necessarily simple. For

example, Kratochwill, Levin, Horner, and Swoboda (2014) reviewed the research on the interpretation of single-subject design graphs, many of which were “simple” A-B designs, and concluded that it was difficult for viewers to reliably visually analyze the graphs in order to determine intervention effectiveness. Difficulties with graph interpretation are not unique to education, or to special education. Research on graph reading in general demonstrates that graph reading is a fairly complex process, and that people easily make mistakes when reading and interpreting graphs (see Friel, Curcio, & Bright, 2001; Glazer, 2011; Shah & Hoeffner, 2002, for reviews).

A term often used in the graph-reading literature to describe people's ability to read and interpret graphs is *graph comprehension* (Friel et al., 2001). *Graph comprehension* is defined as a viewer's ability to derive meaning from a graph, and includes three key components: (1) the ability to extract the data from the graph—that is, to read the data at a surface level; (2) the ability to integrate and interpret the graphed data—that is, to see the relation between the various data components presented on the graph; and (3) the ability to evaluate the data and interpret it within a given context—that is, to make inferences from the data and link the data to “real life” (see Friel et al., 2001, for a review). Curcio (1981) and Friel et al. (2001) refer to these three components of graph comprehension as *reading the data*, *reading between the data*, and *reading beyond the data*, and argue that the components are hierarchical in nature, with reading the data being the simplest, and reading beyond the data the most complex skill.

Curcio's (1981) and Friel et al.'s (2001) framework of ten has been used in graph-comprehension research (e.g., Boote, 2014; Galesic & Garcia-Retamero, 2011; Kim, Lombardino, Cowles, & Altmann, 2014). Applying this framework to CBM, comprehension of CBM progress graphs can be conceptualized as the ability to (1) *read the data*—that is, describe the scores and growth/slope lines on the graph (e.g., “At week 5 the student had a score of 20 correct maze choices,” or “The slope line for phase 3 increased at a rate of .25 choices per week”); (2) *read between the data*—that is, interpret the relations between various data components such as the slope and goal lines (e.g., “The slope line is less steep than the goal line, so growth is less than expected”); and (3) *read beyond the data*—that is, link the data to the instructional context (e.g., “The student is not growing at the expected rate, thus a change in instruction is needed”). We make use of Curcio's and Friel et al.'s framework for our research on CBM graph-comprehension; however, rather than use the generic terms of *reading*, *reading between*, and *reading beyond the data*, we use terms specific to CBM graph-reading, namely *reading*, *interpreting*, and *linking CBM data to instruction*.

Factors Influencing Graph Comprehension

One consistent finding to emerge from the graph-comprehension research is that general graph-literacy can affect the viewer's comprehension of a particular graph (e.g., Glazer, 2011). Graph literacy refers to the viewer's knowledge about graphs (Shah & Hoeffner, 2002). For example, Xi (2010) found that viewers who were more familiar with graphs (i.e., had a higher level of graph literacy) described line graphs in a more organized fashion, and were more complete, accurate, and sophisticated in their graph descriptions, than viewers who were less familiar with graphs. In this study, we examine the role of general graph-literacy in teachers' comprehension of CBM graphs. We measure graph literacy via both self-report and a graph-reading skills test, approaches that have been used in other studies of graph comprehension (e.g., Galesic & Garcia-Retamero, 2011; Xi, 2010).

A second factor that has been found to influence graph comprehension is content knowledge (e.g., Friel et al., 2001; Glazer, 2011). Content knowledge refers to the viewer's background knowledge about the information being graphed (Friel et al., 2001). For example, Shah (2002) found that when viewers were more familiar with the graph content (of line graphs), they were more likely to extract information on trends in the data than when they were less familiar with the graph content.

The effects of content knowledge often have been studied by comparing the graph comprehension of participants with more or less content knowledge (experts versus non-experts; see Freedman & Shah, 2002, for examples of such studies). With regard to CBM, defining “content” knowledge is somewhat of a challenge because content knowledge might be defined as general knowledge about education, general knowledge about educational progress-monitoring, specific knowledge about CBM progress-monitoring, or knowledge related to the individual student being monitored.

In this study, we examine the influence of various sources of content knowledge on CBM graph-comprehension in two ways. First, using standard (researcher-made) CBM graphs, we compare teachers' graph comprehension to that of three groups of experts: general graph-reading experts, education graph-reading experts, and CBM graph-reading experts. Second, to examine the influence of knowledge related to the individual student being monitored, we compare teachers' comprehension of standard (researcher-made) graphs to their comprehension of student graphs from two students with reading difficulties from their own class.

What Should Be Expected of Teachers?

One challenge we faced in conducting this research was knowing what to expect from the teachers with regard to CBM graph-comprehension. Research on CBM graph-comprehension is fairly new, and thus there were few standards against which to compare teachers' performance. Including data from the experts provided us with a standard against which to interpret the teachers' data. This approach also was taken in a study by Wagner, Hammerschmidt-Snidarich, Espin, Seifert, and McMaster (this issue), who examined preservice teachers' comprehension of CBM graphs, and compared those data from the preservice teachers to that of three “gold-standard” CBM experts. Wagner et al. used the term “gold-standard” to emphasize that data from the experts set a standard against which to compare data from the preservice teachers. In this study, we refer to the CBM expert data reported in Wagner et al. to provide a framework for interpreting the data from our inservice teachers. In addition, we extend the Wagner et al. study by including additional variables that were not examined in the original study, and by including data from general graph-reading experts and education graph-reading experts.

Purpose of the Study

To summarize, this study is a replication and extension of Wagner et al.'s (this issue) study on comprehension of CBM graphs. This study is an exploratory, descriptive study, with the purpose of examining inservice teachers' comprehension of CBM graphs, and exploring the influence of factors that might affect that comprehension.

To examine CBM graph-comprehension, we employ a think-aloud strategy, and collect data from teachers on both standard and student CBM graphs. For the standard graphs, we also present data from three types of gold-standard experts. In addition, we examine the relation between teachers' graph-literacy and CBM graph-comprehension.

METHOD

Participants

Teachers

Teacher participants were 23 Dutch elementary- and secondary-school teachers (19 female, 4 male; $M_{age} = 42.39$,

$SD = 11.91$) from 13 general and special education schools who were recruited via convenience sampling. All participants had completed a teacher education program and earned a Bachelor of Education. In addition, 5 teachers had completed or were completing a university-level Bachelor or Master of Science program.¹

Teachers reported that they had had, on average, 4.65 years ($SD = 1.27$, range 2–7 years) of mathematics education during their secondary-school education. Five teachers also had completed one or more (range 1–4) courses in statistics as part of their post-secondary education. Elementary-school teacher participants ($n = 19$) taught at the 5th- and 6th-grade level, and had on average 16.74 years ($SD = 10.31$) of teaching experience. Secondary-school teacher participants ($n = 4$) taught Dutch at the 7th- and 8th-grade level, and had on average 13.25 years ($SD = 9.43$) of teaching experience. All teachers worked with students with reading difficulties in their classes.

Teachers completed a short background questionnaire to assess their familiarity and/or experience with progress monitoring in general, and with CBM progress-monitoring in particular. CBM progress-monitoring is relatively new in the Netherlands, but the concept of progress-monitoring is not. At the elementary level, schools are required to monitor the progress of all students in the school. Most schools use a nationally-normed standardized test to monitor student progress, and students typically are tested annually or bi-annually. Both individual and class-wide data are provided to teachers in the forms of graphs and tables. At the secondary level, progress monitoring is not required, but schools are strongly encouraged to do so. A national standardized test is also available for secondary schools that wish to implement progress monitoring.

Twenty teachers in our sample reported that their schools implemented a progress-monitoring system, and 18 of those teachers reported that they used the data and progress graphs generated by the system. Those teachers reported that they used data to examine student progress, to place students into instructional groups, or to report on student progress to parents. Only 5 of the 23 teachers reported that they had ever heard of CBM progress-monitoring—two via University coursework and one via participation in a study in which teachers collected CBM data from students but did not graph or use the data. None of the teachers had ever used CBM to monitor the progress of students in their classes and evaluate instructional effectiveness.

“Gold-Standard” Graph-Reading Experts

Expert participants were seven “gold-standard” graph-reading experts (3 female, 4 male). Three types of “gold-standard” experts were included: *General-graph Experts*, *Education-graph Experts*, and *CBM-graph Experts*. *General-graph Experts* ($n = 2$, $M_{age} = 35.00$) were assistant professors in Statistics, and were selected because of their training and experience in reading numerical and statistical graphs. Both experts had a master’s degree in Psychology and a Ph.D. in Psychology/Statistics. The *General-graph Experts* had on average 10.50 years of experience teaching statistics; one had

taught 6 different statistics courses, and the other 9. Courses taught by the experts included *Introduction to Statistics & Research Methods*, *Test Theory & Scale Development*, and *Applied Multivariate Data-analysis*.

Education-graph Experts ($n = 2$, $M_{age} = 33.86$) were employees (one full time, the other a consultant) of a company responsible for the development and use of national standardized assessments in the Netherlands (similar to the ETS in the United States). These experts were selected because of their training and experience in reading educational progress graphs. Both experts had a master’s degree in Psychology. One had a Ph.D. in Education and Child Studies, the other a Ph.D. in Psychology/Statistics. (At the time of the study, this second expert was also an assistant professor in Education.) The *Education-graph Experts* had worked on average 7.50 years for the assessment company, and were responsible for the development of language and math items and tests. Both experts had given presentations about interpretation and use of national standardized assessment data to (future) educational professionals.

The *CBM-graph Experts* ($n = 3$, $M_{age} = 66$) were University professors in Special Education, and were selected because of their training and experience in reading CBM graphs. All three *CBM-graph Experts* had Ph.D.’s in Educational Psychology, and were involved in the original development of CBM. They all had at least 100 publications on CBM and had given more than 50 courses or training workshops related to CBM, and reported that they had interpreted more than 100 CBM graphs.

As is clear from the descriptions above, expertise was established primarily on the basis of background and experience; however, we also collected data on experts’ graph literacy. These data are reported at the beginning of the results section.

Procedures

We employed a think-aloud strategy to collect data on participants’ CBM graph-comprehension. Think-aloud data for the teachers and *General-graph* and *Education-graph Experts* were collected as a part of this study. Think-aloud data for the *CBM-graph Experts* had been collected as a part of the Wagner et al. study (see this issue). We used similar procedures as those used by Wagner et al., with the exception that we collected eye-movement data from our participants while they described the graphs. (We report on only the think-aloud data in this article.)

We extended the Wagner et al. (this issue) study by including additional variables on CBM graph-comprehension. For variables common to both the Wagner et al. study and this study, we refer to the Wagner et al. data. For variables unique to this study, we recoded the *CBM-graph Experts’* think-aloud data.²

Teachers completed think-alouds for both standard and student graphs. To create the student graphs, teachers collected weekly progress-monitoring data for two students with reading difficulties over a period of 10 to 12 weeks. Data were collected via an online progress-monitoring system that automatically timed the measures, and scored and graphed the data.³ The CBM measure used for progress monitoring was

maze-selection. A maze is a text in which every seventh word is deleted and replaced by three alternatives. Students read the text silently for two minutes, selecting words at each deletion point. The number of correct and incorrect choices are scored and graphed. Scores from the maze have been found to be reliable and valid indicators of students' performance and progress in reading (Espin, Wallace, Lembke, Campbell, & Long, 2010; Shin, Deno, & Espin, 2000; Wayman et al., 2007).

After collecting progress data for 10–12 weeks, teachers rated their graph-interpretation experience and completed a Graph-Reading Skills Test online. Teachers then completed think-alouds for two standard and two student CBM graphs. General-graph and Education-graph Experts rated their graph-interpretation experience, and completed the Graph-reading Skills Test and then the think-alouds for the standard graphs. The CBM-graph Experts rated their graph-interpretation experience and completed the Graph-reading Skills Test as part of this study.

Think-alouds were conducted on an individual basis. Participants were shown a sample CBM graph in reading, were provided with a description of the graph, and then completed a think-aloud for each standard CBM graph. The order in which the graphs were presented was counterbalanced (AB versus BA). Teachers (only) then completed think-alouds for their students' graphs. Prior to completing the think-alouds for student graphs, teachers were given a short set of instructions describing the differences in layout between the standard and student graphs. Data for the teachers were collected at their school. Data for the experts were collected at their place of work.

Materials

Standard Graphs

The standard (researcher-made) CBM graphs used in this study were slightly modified versions of graphs used in the Wagner et al. study (this issue). For this study, the y-axis represented scores on maze-selection rather than reading aloud because teachers were using the maze to collect progress data from their own students. Although the graphs had a different scale, the data points and data patterns for the graphs used in this study and in the Wagner et al. study were the same.

Standard graphs depicted fictitious student progress data across five phases of instruction across a school year (see Figure 1 for a sample standard graph). The graphs included baseline and peer data, a goal line, and, within each phase, data points and slope lines. The graphs were in black and white and included a legend defining the graph symbols. The format of the sample graph, which was used to provide instructions to participants, was identical to that of the two standard graphs but the data differed.

Student Graphs

Student graphs were created via the progress-monitoring system used to collect progress data. The student graphs had a different format than that of the standard graphs (see Figure 2

for a sample student graph). Student data were collected for a period of only 10–12 weeks, thus the graphs depicted progress for only one instructional phase. In addition, the graphs did not display peer data, and they were in color.

Measures: Graph Literacy

Participants' graph literacy was measured via a self-report question on graph-interpretation experience and a Graph-reading Skills Test.

Self-Report Question Graph-Interpretation Experience

Participants were asked to rate their experience with interpreting graphs and diagrams on a four-point scale ranging from *very little* (1) to *very much* (4).

Graph-Reading Skills Test

The Graph-reading Skills Test was a revised version of the Graph Literacy Scale developed by Galesic and Garcia-Retamero (2011). The original scale was used to assess health-related graph literacy in Germany and the United States (U.S.), and consisted of 8 graphs (bar graphs, line graphs, a pie chart, and an icon array) and 13 questions. Questions were designed to represent Curcio's (1981) three components of graph comprehension (i.e., reading the data, reading between the data, and reading beyond the data). In the Galesic and Garcia-Retamero (2011) study, the scale was administered to nationally representative samples of 495 German and 492 U.S. participants, ages 25 to 69. The scale was found to have reasonable psychometric properties: Cronbach's alpha was .74 for the German version and .79 for the English version, and the total score on the scale correlated significantly with participants' educational level ($r = .29$ for Germany; $r = .54$ for the U.S.) and numeracy skills ($r = .32$ for Germany; $r = .50$ for the U.S.), and with graph-reading items from other measures ($r = .32$ for Germany; $r = .50$ for the U.S.).

We modified the items of the Graph Literacy Scale to fit the purpose of the current study.⁴ Items were changed to reflect educational rather than health-related topics, and were translated from English into Dutch. The first author, who was fluent in both English and Dutch, translated the items. The second and the third authors, who also were fluent in both English and Dutch, reviewed the translation and provided feedback. Then the test was administered to 10 master's students in Education and Child Studies who provided feedback on the items. Items were revised slightly on the basis of this feedback. In addition, an item was added that included a graph that was similar to the progress graphs commonly used in the Netherlands. As a final step, the Graph-reading Skills Test was translated back to English by the researchers so that the CBM-graph Experts could complete the measure.

Participants' scores on the Graph-reading Skills Test were the number of items answered correctly, with a maximum possible score of 14. Cronbach's alpha for the test was .81.

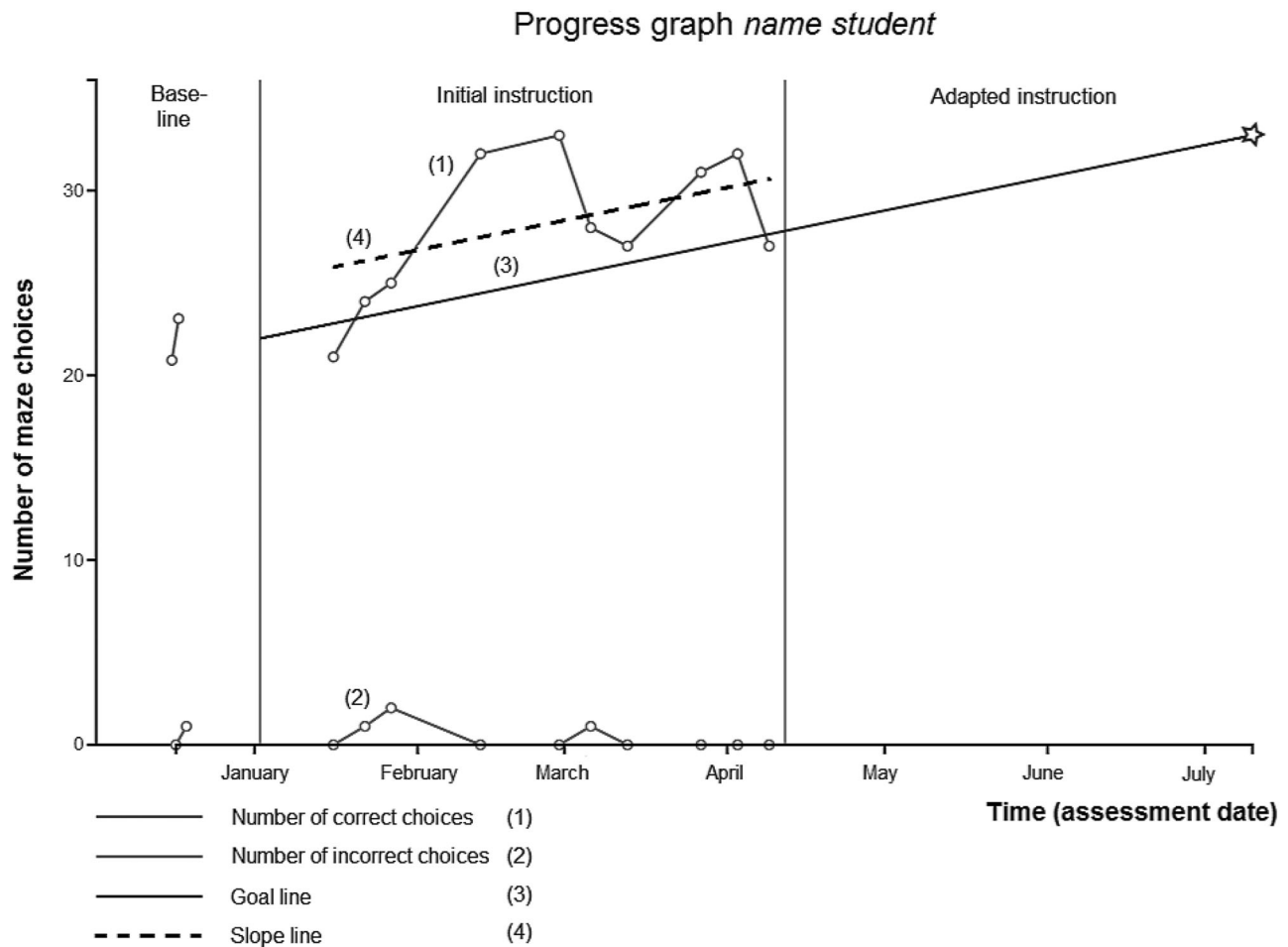


FIGURE 2 Sample of a student CBM graph. Graphs were presented to teachers in Dutch and in color. Numbers were added to this graph to be read in black and white. In the original graphs, correct choices were in green, incorrect choices in blue, the goal line in red, and the slope line in black.

Measures: CBM Graph-Comprehension

Participants' CBM graph-comprehension was assessed via a think-aloud methodology. In a think-aloud methodology, participants are asked to verbalize their thoughts while completing a task (Ericsson & Simon, 1993).

Think-Alouds

Our participants were asked to "think out loud" as they described CBM graphs. They were provided with the following instructions: "Describe the graph and think-out loud while you are looking at the graph. Tell me what you see and what you think. Tell me also where you are looking at and why you are looking at that."

Prior to completing the think-alouds, participants were shown a sample of, and provided with a description of, a CBM graph.⁵ Participants were told that the graph displayed the reading progress of one student across a school year, and that the data on the graph represented correct and incorrect responses on 2-minute reading probes administered weekly to students. The researcher then pointed to and de-

scribed each element of the graph (see Appendix for this description).

Think-alouds were audiotaped and transcribed. Each transcription was checked by a second person, who listened to the tape while reading the transcription, and made corrections if necessary. Disagreements, such as unclear utterances, were resolved by the first author.

Coding Procedures for Standard Graphs

Think-alouds were coded based on the three components of Curcio's and Friel et al.'s framework for graph comprehension (Curcio, 1981; Friel et al., 2001). Recall that we used CBM-specific terms for reading, reading between, and reading beyond the data, namely reading, interpreting, and linking the data to instruction. Coding was done by the first author and by research assistants trained by the first and second author. Coders were trained in five training sessions. Each training session focused on a different aspect of the coding procedure, and included an explanation of the procedure, opportunities for practice, and a reliability check.

Coders had to be 80 percent reliable before they could begin coding.

All data were double coded by the first author and a research assistant. Disagreements in coding were discussed and resolved. Intercoder agreement was calculated separately for each aspect of the coding. To calculate agreement, every third think-aloud was randomly selected, and coding agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements, multiplied by 100.

Two rounds of coding were done. The first focused on participants' ability to read the data. The second focused on participants' ability to interpret the data and link it to instruction.

Round 1: Coding for Reading the Data

Procedures for coding for reading the data were based on procedures developed in previous research (see Espin, Wayman, Deno, McMaster, and De Rooij, this issue; Wagner et al., this issue). Prior to coding, the think-alouds were parsed into idea units (defined as a statement that expressed one idea), and were assigned content labels corresponding to the element of the graph to which they referred, using the definitions from Espin et al. Graph elements included *Framing* (i.e., describing the graph-set up and meaning of the scores or measures used); *baseline* (i.e., describing the beginning level of performance of the student and/or peers); *goal setting* (i.e., describing the goal line and/or long- or short term goals); *instructional phases 0, 1, 2, 3, and 4* (i.e., describing scores, progress, or variability within a specific phase); and *goal achievement* (i.e., describing whether the student achieved the goal). Statements that referred to general student progress (across phases) rather than to progress within a phase were assigned a general progress label. Statements that did not refer to graph content (e.g., comparing one graph to

the other) and evaluative statements about the information on the graph (e.g., wondering why the student had reading problems) were assigned a label of "other." Statements that were irrelevant to the content of the graph (e.g., asking if they were speaking loud enough) were not coded. To illustrate the content label coding, a sample of a coded think-aloud is provided in Table 1. Intercoder agreement for content label coding was 79.70 percent.

After each idea unit was assigned a content label, the think-alouds were coded for three different aspects of reading the data: Accuracy, completeness, and sequential coherence.

Accuracy was the extent to which the statements in the think-aloud were correct. Incorrect statements were those that clearly conflicted with the data presented in the graph—for example, if a participant stated that a student was making progress, but the slope line on the graph was negative. Accuracy was reported as a percentage score, and was calculated by dividing the number of idea units coded as accurate by the total number of idea units. Higher scores reflected a more accurate think-aloud. Intercoder agreement for accuracy was 95.27 percent.

Completeness was the extent to which the think-aloud included mention of nine graph elements: Framing, baseline, goal-setting, phases 0, 1, 2, 3, and 4, and goal achievement. One point was assigned for each element mentioned. The completeness score thus ranged from 0 to 9, with a higher score reflecting a more complete think-aloud. Intercoder agreement for completeness was 100 percent.

Sequential coherence was the extent to which participants described the nine graph elements (see Completeness) in a coherent and logical manner. The concept of sequential coherence was developed in an earlier study (see Espin et al., this issue), and reflected the sequential steps one would take to create and use CBM graphs for evaluation of student growth and instructional effectiveness. The ideal sequence is one in which participants describe the graph elements in the following order: From the set-up of the graph

TABLE 1
Sample of a Coded Think-Aloud

<i>Transcription of the Think-Aloud</i>	<i>Content Label</i>
1. This is the graph of a 6th-grade student.	FR
2. First I look at the current level of performance of the student to find out how this student performs in comparison to peers.	BL
3. Then I look at the long term goal that has been set for this student. The goal for the student is to be at the current level of his/her peers.	GS
4. During initial instruction, <u>some of the student's scores are above the goal,</u> ² but the slope is negative: the line decreases.	P0
5. <u>So a change was made</u> ³ to help the student to achieve the goal. <u>This change was effective,</u> ³ <u>the student is heading towards the goal.</u> ²	P1
6. After that another change was made, <u>but this change was less positive.</u> ¹ <u>The student performed less well than in the previous phase.</u> ¹ <u>The student grows somewhat, but at this rate he will not achieve the goal.</u> ²	P2
7. During the next change, intervention 3, we see a small increase. <u>The student's growth is better.</u> ¹	P3
8. <u>The slope line of phase 4 is again very steep, similar to phase 1,</u> ¹ <u>but the scores are higher.</u> ¹ <u>I would thus recommend the instruction of phase 4 for this student.</u> ³	P4
9. This student achieves the goal.	GA

Note. FR = Framing the data; BL = Baseline data, GS = Goal Setting, P0 = Phase 0 (initial instruction) data, P1 = Phase 1 data, P2 = Phase 2 data, P3 = Phase 3 data, P4 = Phase 4 data, GA = Goal Achievement.

¹data-to-data comparison

²data-to-goal comparison

³data-to-instruction link

(framing) to baseline, to goal-setting, to the consecutive instructional phases (P0-P4), to goal achievement. In the original Espin et al. (this issue) study, a higher sequential coherence score was found to relate to higher expert ratings of teacher think-alouds.

To code sequential coherence, the number of adjacent think-aloud statements that followed the “ideal” sequence were coded, for example, from framing to baseline (1 ideal sequence), baseline to goal-setting (1 ideal sequence), goal-setting to Phase 0, initial instruction (1 ideal sequence), and so forth. If a participant described framing and then Phase 4 instruction, it was not scored as an “ideal” sequence. Statements coded as “other” were ignored in the sequential coherence coding. Sequential coherence was reported as a percentage score, and was calculated by dividing the number of sequences in the ideal order by the total number of sequences. Sequences that included a general progress statement were excluded from this calculation. Higher sequential coherence scores reflected a more coherent think-aloud. Inter-coder agreement for sequential coherence was 94.44 percent.

Round 2: Coding for Interpreting and Linking the Data to Instruction

Within the second round of coding, think-alouds were coded for two aspects of interpreting the data. We refer to these aspects as data-to-data and data-to-goal comparisons. Data also were coded for one aspect of linking the data to instruction. We refer to this aspect as data-to-instruction links.⁶ (In the sample of the coded think-aloud in Table 1, examples of these comparisons and links are underlined.)

Data-to-data comparisons were counted when participants compared data in one instructional phase to data in another instructional phase. For example, the participant might comment on differences in student growth across phases.

Data-to-goal comparisons were counted when participants compared student performance or progress data to the goal line or the end-of-year goal. For example, the participant might comment on whether the data indicated that the student was on track for achieving the goal. Data-to-goal comparisons could involve comparisons with regard to level (e.g., “Student performance is below the goal line”) or rate (e.g., “The student was progressing at the expected rate”).

Data-to-instruction links were counted when participants linked the data in the graph to the student’s reading instruction. For example, the participant might comment on the fact that a positive slope indicated that the instruction was effective. Inter-coder agreement for this round of coding was 80.09 percent.

Coding Procedures for Student Graphs

Student graphs differed from teacher to teacher because teachers viewed and described graphs from their own students. Recall that student graphs included only one instructional phase; thus, think-alouds could not be coded for completeness, accuracy, or sequential coherence, as was done for the standard graphs. However, they could be coded for

interpreting and linking the data to instruction. With regard to interpreting the data, only data-to-goal comparisons were coded. (There was only one instructional phase, so data-to-data comparisons could not be made.) In sum, data-to-goal comparisons and data-to-instruction links were coded for the student graphs. Inter-coder agreement for coding of students graphs was 90.36 percent.

RESULTS

We first report descriptive statistics on the graph-literacy measures for teachers and experts. We then report on the think-aloud data for the standard graphs for teachers and experts, and then on the student graphs for teachers only. Finally, we report on the relation between teachers’ graph literacy and CBM graph-comprehension.

Participants’ Graph Literacy

Both teachers and experts completed the graph-literacy measures. An independent samples t-test and a Mann-Whitney U-test were conducted to compare teachers’ and experts’ graph-literacy scores. Scores for self-reported graph-interpretation experience were significantly lower for the teachers ($M = 2.83$, $SD = 0.72$, range 2–4) than for the experts ($M = 3.71$, $SD = 0.49$, range 3–4), $t(28) = -3.05$, $p < .01$, $d = 1.15$. Obtained scores on the Graph-reading Skills Test were lower for teachers ($M = 11.57$, $SD = 2.69$, $Mdn = 12$, range 3–14) than for the experts ($M = 12.71$, $SD = 1.38$, $Mdn = 13$, range 11–14), but the difference was not significant, $U = 56.50$, $z = -0.93$, $p > .05$. There was a ceiling effect on the Graph-reading Skills Test (details reported later).

CBM Graph-Comprehension: Standard Graphs

Our first set of analyses focused on teachers’ comprehension of the two standard graphs. Average scores across the think-alouds for the two graphs were used in all analyses. Teachers’ think-alouds for the standard graphs varied in length from 45.50 to 470.50 words ($M = 204.02$, $SD = 125.29$) and in the number of idea units from 3 to 24.50 idea units ($M = 11.09$, $SD = 5.32$). Think-alouds for the General-graph, Education-graph, and CBM-graph Experts were longer than for the teachers, with an average of 398, 278.25, and 556.50 words, and 11.50, 16.25, and 17.67 idea units, respectively.

Reading the Data

Descriptive statistics for accuracy, completeness, and sequential coherence (the three aspects of reading the data) are reported in Table 2. Teachers were fairly accurate in their think-alouds, with an average accuracy of 98 percent (range 87.50–100). Only 6 of the 23 teachers made any inaccurate statements. Accuracy scores for teachers were similar to

TABLE 2
Descriptive Statistics on Participants' CBM Graph-Comprehension Scores for Standard Graphs

<i>CBM Graph-Comprehension Score</i>	<i>Teachers (n = 23)</i> <i>M (SD)</i>	<i>General-Graph Experts (n = 2)</i> <i>M</i>	<i>Education-Graph Experts (n = 2)</i> <i>M</i>	<i>CBM-Graph Experts (n = 3)</i> <i>M</i>
Accuracy (percentage)	97.53 (4.47)	95.56	100	100
Completeness (score out of 9)	5.72 (2.37)	4.75	7.75	8.33
Sequential coherence (percentage)	51.71 (33.17)	22.98	59.72	85
Data-to-data comparisons (number)	1.67 (1.47)	4	4	4.83
Data-to-goal comparisons (number)	1.72 (1.49)	0.50	1.25	4.17
Data-to-instruction links (number)	0.98 (1.26)	1	2.75	5

Note. Accuracy, completeness, and sequential coherence scores reflect participants' ability to read CBM data; the number of data-to-data and goal-comparisons reflect participants' ability to interpret CBM data; and the number of data-to-instruction links reflects participants' ability to link CBM data to instruction.

those of the experts, whose average accuracy ranged from 96 percent to 100 percent.

Teachers were moderately complete in their think-alouds, mentioning on average 6 of 9 possible graph elements in their think-alouds, with scores ranging from 1 to 9. Goal achievement and data from instructional phase 1 were described most often, while framing, baseline data, and goal setting were described least often. Teachers were more complete than the General-graph Experts, who mentioned on average 5 out of the 9 graph elements, but less complete than the Education-graph and CBM-graph Experts, who both mentioned on average 8 graph elements.

With regard to sequential coherence, results revealed that teachers were moderately coherent, with an average sequential coherence of 52 percent. However, variability was high, with scores ranging from 0 percent (for 5 teachers) to 100 percent (for 2 teachers). Teachers were more coherent in their think-alouds than the General-graph Experts, who had average coherence scores of 23 percent, but less coherent than the Education-graph and CBM-graph Experts, who had average coherence scores of 60 percent and 85 percent, respectively.

Interpreting and Linking the data to Instruction

With regard to *interpreting the data*, teachers made on average 2 data-to-data and 2 data-to-goal comparisons (see Table 2), with a *range* of 0 to 5.50 comparisons for each. Twenty teachers made at least 1 data-to-data comparison and 19 teachers made at least 1 data-to-goal comparison. Teachers made fewer data-to-data comparisons (with an average of 2) than the General-graph, Education-graph, and CBM-graph Experts, who made an average of 4, 4, and 5 data-to-data comparisons, respectively. Teachers made more data-to-goal comparisons (with an average of 2) than the General-graph and Education-graph Experts, who both made an average of 0.5 to 1 data-to-goal comparison, but fewer than the CBM-graph Experts, who made an average of 4 data-to-goal comparisons.

With regard to *linking the data to instruction*, results revealed that teachers made on average only 1 data-to-instruction link (see Table 2), with a *range* of 0 to 4 links. Only 11 teachers made at least 1 data-to-instruction link in their think-alouds. Teachers made the same number of links as the General-graph Experts, who also made 1 link, but

fewer than the Education-graph and CBM-graph Experts, who made 3 and 5 links, respectively.

CBM Graph-Comprehension: Student graphs

Our second set of analyses focused on teachers' comprehension of the student graphs. Recall that the student graphs were coded only for interpreting (and only for data-to-goal comparisons) and for linking data to instruction. Average scores across the think-alouds of the two student graphs were used in the analyses.⁷

Interpreting and Linking the Data to Instruction

With regard to *interpreting the data*, teachers made 1.22 ($SD = 0.85$) data-to-goal comparisons, with a *range* from 0 to 4. Twenty teachers made at least 1 data-to-goal comparison. With regard to *linking the data to instruction*, teachers made 0.28 ($SD = 0.58$) data-to-instruction links, with scores ranging from 0 to 2. Only six teachers made at least 1 data-to-instruction link in their think-alouds for the student graphs.

CBM Graph-Comprehension: Standard versus Student Graphs

To compare results across standard and student graphs, the proportion of teachers who made at least one data-to-goal comparison or data-to-instruction link was calculated. The results of two McNemar's tests using a binomial distribution revealed no significant difference in the proportion of teachers who made at least one data-to-goal comparison for the standard graphs (83 percent) and the student graphs (87 percent), $p > .05$, and no significant difference in the proportion of teachers who made at least one data-to-instruction link for the standard graphs (48 percent) and the student graphs (26 percent), $p > .05$.

Relation between Graph Literacy and CBM Graph-Comprehension

Correlational analyses were conducted to examine the relations between teachers' graph literacy, as measured via a self-report question and a Graph-reading Skills Test, and their

TABLE 3
Correlations between Teachers' Graph-literacy Scores and CBM Graph-comprehension Scores for Standard and Student Graphs

Graph-Literacy Measures	CBM Standard Graphs					CBM Student Graphs		
	Accuracy	Completeness	Sequential coherence	Data-to-Data comparisons	Data-to-Goal comparisons	Data-to-Instruction links	Data-to-Goal comparisons	Data-to-Instruction links
Self-report question graph-interpretation experience	-.22	.35	.48*	.25	.65**	.43*	.07	.52**
Graph-reading Skills Test								
All items	.09	-.18	-.02	.14	-.22	-.04	-.12	-.34
5 discriminating items	.24	.23	-.03	.12	.28	.29	-.25	.46*

Note. $N = 21$. Correlations in *italics* are Pearson correlations; the others are Kendalls' tau correlations.

* $p < .05$.

** $p < .01$.

comprehension of the standard and student CBM graphs, as measured via think-alouds.

Prior to the correlational analyses, accuracy and sequential coherence scores were transformed with arcsine-square root transformation because proportion variables do not have a normal distribution (Cohen & Cohen, 1983; Osborne, 2009). For the standard graphs, teachers' (transformed) think-aloud scores (i.e., accuracy, completeness, sequential coherence, number of data-to-data and data-to-goal comparisons, and number of data-to-instruction links) were all normally distributed. For the student graphs, teachers' think-aloud scores (i.e., number of data-to-goal comparisons and data-to-instruction links) were not normally distributed. For analyses involving variables that were non-normally distributed, non-parametric tests were used. Kendall's tau (τ) was used rather than Spearman's rho for those non-parametric tests because our sample was small and included tied ranks for the think-aloud scores (Field, 2009). For the analyses involving variables that were normally distributed, Pearson (r) was used. Results of the correlational analyses are reported in Table 3.

Relation between Self-Report Question and CBM graph-comprehension

Teachers' mean score on the self-report question about graph-interpretation experience was 2.83 ($SD = 0.72$, range 2–4), and their scores were normally distributed. Teachers' scores for the self-report question correlated significantly with sequential coherence scores ($r = .48$, $p < .05$) and with the number of data-to-goal comparisons ($r = .65$, $p < .01$) and data-to-instruction links ($r = .43$, $p < .05$) made for the standard graphs. In addition, teachers' scores for self-report question correlated significantly with the number of data-to-instruction links made for the student graphs ($r = .52$, $p < .01$). No other correlations were significant (see Table 3).

Relation between Graph-reading Skills Test and CBM graph-comprehension

Two teachers did not complete the Graph-reading Skills Test, and were thus not included in analyses with scores for this

test. Teachers' mean score on the Graph-reading Skills Test was 11.57 ($SD = 2.69$, range 3–14) out of 14. Teachers' scores for this test were not normally distributed; the distribution of scores was strongly negatively skewed and kurtotic (*standardized skewness* = -4.06 and *standardized kurtosis* = 4.63). A relatively large number of items on the test were answered correctly by all, or nearly all, teachers. Therefore, we conducted an item analysis to identify which items discriminated best, and correlational analyses that included scores for the Graph-reading Skills Test were conducted both with the scores on the total test (14 items) and with scores on the discriminating items only (see Table 3).

To select items that were discriminating, difficulty levels and discrimination indices for each item were calculated. Difficulty levels (p -values) were calculated by dividing the number of teachers who answered the item correctly by the total number of teachers completing the test. Discrimination indices (d -values) were calculated by subtracting the proportion of teachers from the bottom quartile (27 percent, to be exact, $n = 6$) who answered the item correctly from the proportion of teachers in the top quartile (27 percent, $n = 6$) who answered the item correctly (Reynolds & Livingston, 2012). Scores on the test for teachers in the bottom group ranged from 3 to 11, and for teachers in the top group, ranged from 13 to 14 (the maximum score).

Guidelines suggested by Reynolds and Livingston (2012) were used to identify and select the most discriminating items. Items with p -values below .90 and d -values above .30 were selected as discriminating items. Five items met those criteria. Two of those five items represented the reading between the data component of graph comprehension, and the other three items represented the reading beyond the data component.

For teachers' total scores on the Graph-reading Skills Test, the correlations between test scores and CBM graph-comprehension scores ranged from -.34 to .14, and none were significant (see Table 3). For the five discriminating items, the correlations between test scores and CBM graph-comprehension scores ranged from -.25 to .46 and were only significant for the number of data-to-instruction links made for the student graphs ($\tau = .46$, $p < .05$). No other correlations were significant.

DISCUSSION

The purpose of this study was to examine inservice teachers' CBM graph-comprehension, and to explore the influence of factors that might affect that comprehension. We employed a think-aloud strategy, and collected data from teachers on both standard and student CBM progress graphs. For the standard graphs, we also included data from three types of gold-standard graph-reading experts. Finally, we examined the relation between teachers' graph literacy and CBM graph-comprehension.

CBM Graph-Comprehension: Standard Graphs

Reading the Data

Teachers' comprehension of the CBM standard graphs was reasonably good. Teachers' think-alouds were accurate and were moderately complete and coherent. In fact, teachers' performance on these three *reading the data* aspects was similar to, or even better than, that of the General-graph-reading Experts. Especially interesting was the fact that teachers were much better at coherently describing the CBM graphs than were General-graph Experts, illustrating the importance of educational knowledge in the comprehension of CBM graphs. The General-graph Experts had ample experience and skill with graph reading, but not with reading graphs that displayed the progress of students with reading difficulties. The teachers, in contrast, had more educational knowledge than the General-graph Experts, and this knowledge apparently provided teachers with enough information to read and describe the CBM graphs in a more coherent manner than the General-graph Experts. Supporting this explanation is the fact that sequential coherence scores for the teachers were only somewhat lower than scores for the Education-graph Experts. The difficulties experienced by the General-graph Experts were surprising, given the fact that, prior to completing the think-alouds, participants were shown a sample graph and were more or less told what was on the graph. Yet telling a coherent "story" about a CBM graph proved to be a challenge for the General-graph Experts.

Although comparing the teachers' performance on *reading the data* to that of the General-graph Experts paints a positive picture, comparing it to the Education-graph and CBM-graph Experts paints a bleaker picture. In other words, although the teachers' performance was not bad, there was plenty of room for improvement. Compared to both the Education-graph and CBM-graph Experts, teachers were less complete in their think-alouds, and compared to the CBM-graph Experts, they were far less coherent. The differences between teachers and CBM-graph Experts replicate the findings of Wagner et al. (this issue), who found that think-alouds for preservice teachers' were less complete and sequentially coherent than think-alouds for CBM experts.

Comparing the teachers' data from our study to that from Espin et al.'s (this issue) study provides support for the idea that there is room for improvement for the teachers. Espin et al. examined think-alouds of experienced CBM users. Teachers in the Espin et al. study had an average of 12 years of

experience using CBM to monitor student progress, and had generated an average of 160 CBM progress graphs. Teachers in our study were new to CBM, had not received intensive CBM training, and had used CBM for only a short period of time to collect data on student progress. The average sequential coherence for the teachers in the Espin et al. study was 71 percent, compared to 52 percent for teachers in this study, suggesting that with training and experience, teachers' ability to read CBM graphed data improves. However, even the experienced teachers in Espin et al.'s study were not as coherent as the CBM-graph Experts from this/Wagner et al.'s study.

Thus far, we have considered only the teachers' mean scores on the *reading the data* aspects, but it is also informative to consider the variation in scores. The range of teachers' scores for sequential coherence was 0 percent to 100 percent, with five teachers having a sequential coherence score of 0 percent, and for completeness, it was 1 to 9 (graph elements mentioned). Such variation in scores might be expected, given the fact that people differ in their general ability to understand and interpret graphs (Galesic & Garcia-Retamero, 2011), yet this variation demonstrates that some teachers have great difficulty reading CBM graphs. It is perhaps worthwhile to note that even in the Espin et al. (this issue) study, where participants were experienced CBM users, wide variation was seen in sequential coherence scores (from 56 percent for teachers who had lower levels of understanding and interpretation of CBM data, to 89 percent for teachers who had higher levels of understanding and interpretation of CBM data). In sum, it seems fair to say that some teachers are in need of more support than others in learning to read CBM graphs. Given the fact that teachers often are expected to share data from CBM progress graphs with team members and/or parents, it will be important to provide additional or different training for teachers who experience difficulties reading CBM graphs.

Interpreting and Linking the Data to Instruction

Examination of the outcomes for the *interpreting and linking the data* aspects shows a somewhat different pattern of outcomes than for the *reading the data* aspects. Teachers made fewer within-the-data (data-to-data plus data-to-goal) comparisons than all three groups of experts, and fewer data-to-instruction links than the Education-graph and CBM-graph Experts. It would seem that the skills of interpreting and linking the data to instruction are more difficult than reading the data for the teachers. However, if there is "good news," it is that these skills also seem to be relatively difficult for the General-graph and Education-graph Experts. For example, although the teachers made fewer within-the-data comparisons (almost 3.5) than the General-graph and Education-graph Experts (4.5 and 5.25 respectively), all three groups made fewer comparisons than the CBM-graph Experts, who made 9 such comparisons. For linking the data to instruction, although both teachers and General-graph Experts made fewer links (1 each) than the Education-graph Experts (2.75), all three groups made fewer links

than the CBM-graph Experts, who made 5 such links. The pattern of differences among the four groups of participants supports Curcio's and Friel's (Curcio 1981; Friel et al., 2001) contention that reading between and beyond the data is more difficult than reading the data, and suggests that teachers are in need of specific, directed instruction on how to interpret CBM data and on how to link it to instruction.

CBM Graph-Comprehension: Student Graphs

It is possible that teachers' difficulties with interpreting CBM data and linking it to instruction were related to the fact that the standard graphs presented fictitious information that was not directly relevant to the teachers. Thus, we also examined teachers' comprehension of student graphs. We anticipated that teachers would be more likely to make data-to-goal comparisons and data-to-instruction links if graphs were from their own students. With student graphs, teachers could bring to bear specific knowledge about the students and the students' instruction. Our expectations were not, however, supported by the data.

Although the percentage of teachers who made data-to-goal comparisons was similar for the standard and student graphs (83 percent and 87 percent respectively), the percentage of teachers who made data-to-instruction links was not. Even though the difference was not significant, it was fairly large and was not in the expected direction (48 percent for standard graphs, 26 percent for student graphs). The difference might be due to chance alone, or might merely reflect the fact that there were more "phases" of instruction on the standard graphs, and thus, more opportunities for teachers to reflect on the link between the data and instruction in those graphs (although, do recall, we did not compare the raw number of statements, but rather the percentage of teachers who made at least one such link). However, if such a difference were to be replicated in a study with a larger sample size, it might reflect the fact that teachers find it difficult, or even threatening, to evaluate instructional effectiveness as it relates to their own students and their own reading instruction. Regardless of the reasons for the difference, our data suggest that in future research, it would be wise to consider the fact that teachers' CBM graph-comprehension might differ for standard versus student graphs, and that their comprehension of student graphs might be affected by emotional factors.

Relation between Graph Literacy and CBM Graph-Comprehension

As a last step in the analyses, we examined the relation between teachers' graph literacy and their CBM graph-comprehension. In sum, results revealed that teachers' self-reported graph-interpretation experience was related to some, but not all, aspects of their CBM graph-comprehension ability. Teachers' scores on the Graph-reading Skills Test, however, were not related to any aspect of CBM graph-comprehension.

Teachers with higher self-ratings on the graph-interpretation experience question produced more coherent think-alouds, and made more data-to-goal comparisons for the standard graphs, and more data-to-instruction links for both the standard and student graphs, than teachers with lower self-ratings. These results are in line with the results of Xi (2010) who found that participants who were more familiar with graphs—as measured with a self-report measure—provided graph descriptions that were better organized and more sophisticated than participants who were less familiar with graphs.

Although the results related to the self-report measures suggest that graph literacy is important for CBM graph-comprehension, conclusions must be tempered by the fact that no relations were found between scores on the Graph-reading Skills test and CBM graph-comprehension. Results were likely affected by the restricted range of scores caused by the ceiling effect on the Graph-reading Skills Test scores. However, even the analyses conducted with the five most discriminating items resulted in only one statistically significant relation: Teachers with higher scores on the five items made more data-to-instruction links in their think-alouds for student graphs than did teachers with lower scores for those items.

The ceiling effect found on the Graph-reading Skills Test was disappointing. We based the test on one used by Galesic and Garcia-Retamero (2011), who reported promising psychometric properties for the test. However, participants in that study were sampled from the general population. In their German and U.S. samples, respectively 79 percent and 72 percent of the participants were lower educated people (i.e., people who completed high school or less). The teachers in our study were more highly educated, with all of them having completed post-secondary teacher education programs, and some of them having completed or completing a university-level Bachelor or Master of Science program. This difference in education might explain why our teachers scored better on the Graph-reading Skills Test than the participants in the Galesic and Garcia-Retamero (2011) study.

To summarize, our results provide tentative support for the importance of graph literacy on teachers' CBM graph-comprehension. Most important is the fact that for both the self-report measure and the Graph-reading Skills Test (discriminating items only, and student graphs only), a relation was found between graph literacy and teachers' ability to link the data to instruction. At the very least, it seems reasonable to suggest that researchers and trainers must take into consideration teachers' level of graph literacy when providing instruction or conducting research on CBM graph-comprehension.

Limitations and Recommendations for Future Research

A major limitation of this study was that the relation between teachers' CBM graph-comprehension and teachers' use of the data for instructional decision-making, and the resulting

effect on student achievement, was not examined. Such a comparison is the next logical step in this line of research.

A second limitation of the study was the small sample size, which limited the external validity of the results. All participating teachers were from a specific region in the Netherlands. Given the exploratory, descriptive nature of the study, the sample size was appropriate. Nonetheless, it will be important to replicate key findings of this study with a larger, more representative sample, including teachers from different regions and/or countries. When replicating the key findings with a larger sample size, it would be important to examine the potential moderating effects of teacher characteristics on teachers' comprehension of CBM graphs.

A third limitation was the think-aloud approach we chose to use in the study. We used an open-ended think-aloud, and provided general instructions to teachers (and experts), merely saying "describe what you see." It might have been tempting for teachers to ignore parts of the graph that they did not understand. The relatively high accuracy scores combined with the moderate completeness scores support this argument; that is, although what teachers said in their think-alouds was usually accurate, they did not talk about all elements of the graph. An alternative approach would be to ask teachers to describe the graph as if they were talking to a student's parent. Such an approach might encourage teachers to attend to all graph elements, and might make the task more realistic for teachers.

A final limitation of the study was that student graphs had a different lay-out and depicted less information than the standard graphs, and that they differed from teacher to teacher. These differences made it somewhat challenging to compare teachers' performance on the standard and student graphs, and to compare performance on the student graphs across teachers. Although including student graphs introduced challenges, it could also be viewed as a major strength of the study because the student graphs presented teachers with a more authentic situation than the standard graphs.

Final Thoughts: Do Teachers Need to Comprehend CBM Graphs?

As a final point in the discussion, we wish to raise the issue of whether teachers actually *need* to comprehend CBM graphs in order to effectively use the graphed data to make instructional decisions. In other words, how important is it for teachers to read, interpret, and link CBM data to instruction, given the fact that computer-based progress-monitoring programs can provide teachers with recommendations or prompts to raise the goal or to change instruction?

This question is an empirical question that must be answered in future research, but it is worthwhile to reflect on the question here. We would argue that, even with computer-generated decision-making supports, it is necessary for teachers to be proficient at reading, interpreting, and linking CBM data to instruction. There are several reasons to believe that computer technology alone will not be enough to guide teachers' instructional decision-making.

First, as mentioned earlier, teachers often must describe and discuss progress data with team members and/or parents.

To explain student progress data effectively, teachers must be able to read and interpret CBM graphs, and be able to link the data to instruction.

Second, CBM data patterns can be ambiguous, potentially leading to two different but "correct" decisions (Espin, Saab, Pat-El, Boender, & Van der Veen, 2016). In such situations, it is imperative that teachers be able to combine CBM progress data with other information about the student to arrive at the "best" decision for that student. Teachers must correctly recognize the ambiguity that is inherent in some data patterns, and be willing to think beyond the recommendation provided by a progress-monitoring program (Deno, 2013).

Finally, computer-generated instructional recommendations alone are not enough to ensure that teachers respond to CBM data. That is, even when prompted to do so via the computer, teachers sometimes do not respond to the data by raising the goal or changing instruction (see Fuchs & Fuchs, 2002; Stecker et al., 2005, for reviews). There are several potential explanations for teachers' non-response. One may be that teachers doubt the meaningfulness and usefulness of data for instructional decision-making (Foegen, Espin, Allinder, & Markell, 2001; Landrum, Cook, Tankersley, & Fitzgerald, 2007), and thus may not trust the instructional recommendations provided by progress-monitoring programs. Improving teachers' comprehension of CBM data might improve their belief in the data, and thereby their willingness to respond to the data. A second explanation may be that teachers have to be involved with the data in order to make effective data-based decisions. Fuchs and Fuchs (Fuchs, 1988; Fuchs & Fuchs, 1989) suggested that computer applications can distance teachers from CBM data, and thereby limit teachers' meaningful interpretation and use of the data. Supporting this idea is a study by Fuchs, Fuchs, and Hamlett (1989) that showed that enhancing teacher involvement with computer-managed CBM led to improvements in timing of goal and instructional changes.

CONCLUSION

In conclusion, the results of this study reveal that comprehension of CBM graphs is not as straight forward as one might assume. Some of our teacher participants were unable to read and describe CBM graphs in a complete and coherent manner, and most of the teachers experienced difficulty with interpreting the data and linking it to instruction, suggesting that teachers need training in or help with reading, interpreting, and linking CBM data to instruction. Such training should be provided to preservice teachers as part of their education, and to inservice teachers in the form of professional development courses. If teachers do not read the data accurately, completely, and coherently, or do not make within-the-data comparisons or data-to-instruction links, it is possible that use of CBM will not lead to improvements in instruction and student achievement. If teachers are expected to base their instructional decisions for students with learning difficulties on data, teachers must be equipped to comprehend graphed student progress data in a way that enables them to engage in data-based decision-making and become successful problem solvers.

ACKNOWLEDGMENTS

We are grateful to the students, teachers, and experts who participated in this study; to Robin van Emden who created the online progress-monitoring system that was used to collect student progress data and to create the student progress graphs used in the study; and to the Master's students and Rianne Feijt who helped with collecting and coding the data for this study.

NOTES

1. The Netherlands has a binary system of higher education. Thus, there are two types of higher education—research-oriented education and professional-oriented education. Research-oriented education is typically offered at research universities, professional-oriented education at universities of applied sciences (EP-Nuffic, 2016). Teacher education falls primarily under the professional-oriented education system.
2. We obtained permission from the authors of the Wagner et al. (this issue) study and from the CBM expert participants to refer to the original data in the present study, and to recode and reanalyze parts of their think-aloud data.
3. Teachers had no access to the student scores or the student graphs; they thus did not see the student graphs before they completed the think-alouds for those graphs.
4. We obtained permission from the developers of the Graph Literacy Scale to modify the items, and to use the modified scale in the present study.
5. Note that the CBM experts in the Wagner et al. (this issue) study were not provided with a description of a sample graph. However, our teachers and experts were new to or had never heard of CBM; thus, we felt that it was necessary to provide an overview of the graphs before asking them to complete the think-alouds.
6. For this round of coding, comparisons or links referring to incorrect choices were not coded. This was done because it is the number of correct choices that reflects progress, not the number of incorrect choices.
7. One teacher described only one student graph. In the analyses, the scores for that one graph were used as the average score.

REFERENCES

- Boote, S. K. (2014). Assessing and understanding line graph interpretations using a scoring rubric of organized cited factors. *Journal of Science Teacher Education, 25*(3), 333–354.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Curcio, F. R. (1981). The effect of prior knowledge, reading and mathematics achievement, and sex on comprehending mathematical relationships expressed in graphs (Doctoral thesis). Retrieved from ProQuest. (8128203)
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*(3), 219–232.
- Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*(3), 184–192.
- Deno, S. L. (2013). Problem-solving assessment. In R. Brown-Chidsey & K. J. Andren (Eds.), *Assessment for intervention: A problem-solving approach* (2nd ed., pp. 10–36). New York: Guilford Press.
- Deno, S. L., & Fuchs, L. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children, 19*(8), 1–16.
- EP-Nuffic. (2016). Higher education system in the Netherlands. Retrieved from <https://www.epnuffic.nl/en/files/documents/factsheet-higher-education-system-in-the-netherlands.pdf>
- Espin, C. A., Saab, N., Pat-El, R., Boender, P., & Van der Veen, J. (2016). Curriculum-based measurement progress data: Effects of graph pattern on ease of interpretation. Unpublished manuscript.
- Espin, C. A., Wallace, T., Lembke, E., Campbell, H., & Long, J. (2010). Creating a progress-monitoring system in reading for middle-school students: Tracking progress toward meeting high-stakes standards. *Learning Disabilities Research & Practice, 25*(2), 60–75.
- Espin, C. A., Wayman, M. M., Deno, S. L., McMaster, K. L., & De Rooij, M. (2017). Data-based decision-making: Developing a method for capturing teachers' understanding of CBM graphs. *Learning Disabilities Research & Practice, 32*(1), 8–21.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis. Verbal reports as data* (revised ed.). Cambridge, MA: The MIT Press.
- Field, A. (2009). *Discovering statistics using SPSS*. Los Angeles: SAGE Publications Ltd.
- Foegen, A., Espin, C. A., Allinder, R. M., & Markell, M. (2001). Translating research into practice: Preservice teachers' beliefs about curriculum-based measurement. *The Journal of Special Education, 34*, 226–236.
- Foegen, A., Jiban, C., & Deno, S. L. (2007). Progress monitoring measures in mathematics: A review of the literature. *The Journal of Special Education, 41*(2), 121–139.
- Freedman, E. G., & Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In M. Hegarty, B. Meyer, & N. Hari Narayanan (Eds.), *Diagrammatic representation and inference* (pp. 18–30). Berlin: Springer-Verlag.
- Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*(2), 124–158.
- Fuchs, L. S. (1988). Effects of computer-managed instruction on teachers' implementation of systematic monitoring programs and student achievement. *Journal of Educational Research, 81*(5), 294–304.
- Fuchs, L. S., & Fuchs, D. (1989). Enhancing curriculum-based measurement through computer applications: Review of research and practice. *School Psychology Review, 18*(3), 317–327.
- Fuchs, L. S., & Fuchs, D. (2002). Computer applications to curriculum-based measurement. *Special Services in the Schools, 17*(1-2), 1–14.
- Fuchs, L. S., Fuchs, D., & Hamlett, C. H. (1989). Computers and curriculum-based measurement: Effects of teacher feedback systems. *School Psychology Review, 18*(1), 112–125.
- Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy: A cross-cultural comparison. *Medical Decision Making, 31*(3), 444–457.
- Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in Science Education, 47*(2), 183–210.
- Kim, S., Lombardino, L. J., Cowles, W., & Altmann, L. J. (2014). Investigating graph comprehension in students with dyslexia: An eye tracking study. *Research in Developmental Disabilities, 35*, 1609–1622.
- Kratochwill, T. R., Levin, J. R., Horner, R. H., & Swoboda, C. M. (2014). Visual analysis of single-case intervention research: Conceptual and methodological issues. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Statistical and methodological advances* (pp. 91–125). Washington, D.C.: American Psychological Association.
- Landrum, T. J., Cook, B. G., Tankersley, M., & Fitzgerald, S. (2007). Teacher perceptions of the usability of intervention information from personal versus data-based sources. *Education and Treatment of Children, 30*(4), 27–42.
- McMaster, K., & Espin, C. A. (2007). Technical features of curriculum-based measurement in writing: A literature review. *The Journal of Special Education, 41*(2), 68–84.

- Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research & Evaluation*, 8(6), 1–8. Retrieved from <http://PAREonline.net/getvn.asp?v=8&n=6>
- Reynolds, C. R., & Livingston, R. B. (2012). *Mastering modern psychological testing, Leiden version*. (Custom edition). Upper Saddle River, NJ: Pearson.
- Shah, P. (2002). Graph comprehension: The role of format, content and individual differences. In M. Anderson, B. Meyer, & P. Oliver (Eds.), *Diagrammatic Representation and Reasoning* (pp. 173–185). London: Springer-Verlag.
- Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, 14(1), 47–69.
- Shin, J., Deno, S. L., & Espin, C. A. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education*, 34(3), 164–172.
- Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools*, 42(8), 795–819.
- Wagner, D. L., Hammerschmidt-Snidarich, S., Espin, C. A., Seifert, K., & McMaster, K. L. (2017). Pre-service teachers' interpretation of CBM progress monitoring data. *Learning Disabilities Research & Practice*, 32(1), 22–31.
- Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, 41(2), 85–120.
- Xi, X. (2010). Aspects of performance on line graph description tasks: Influenced by graph familiarity and different task features. *Language Testing*, 27(1), 73–100.
- Zigmond, N. (1997). Educating students with disabilities: The future of special education. In J. W. Lloyd, E. J. Kamenui, & D. Chard (Eds.), *Issues in educating students with disabilities* (pp. 377–390). Hillsdale, NJ: Erlbaum Associates.
- Zigmond, N. (2003). Where should students with disabilities receive special education service? Is one place better than another? *The Journal of Special Education*, 37(3), 193–199.

Appendix

Description of the Sample CBM Graph

In the graph you see the baseline level of the student (*point to baseline data*). The baseline phase shows the student's level of performance compared to peers at the start of the school year (*point to peer data*). The baseline data are used to set a long-range goal for the student for the end of the school year (*point to goal line and end point of goal line*). In the graph you can see the initial instructional phase for this student (*point to first phase*), followed by four phases representing changes to that instruction (*point to phases*). For each phase, you can see the student's response to the reading instruction in that phase (*point to slope lines*).

About the Authors

Roxette M. van den Bosch, MSc, is a Ph.D. student of the Department of Education and Child Studies (specialization Learning Disabilities) at Leiden University, The Netherlands. Her research interests include data-based decision-making, Curriculum-based Measurement, and teachers' graph-comprehension and data-use skills.

Christine A. Espin, Ph.D., is a Professor in the Department of Education and Child Studies (Special Education / Learning Disabilities) at Leiden University, The Netherlands. Her research interests include teachers' data-based decision-making, Curriculum-based Measurement in reading, writing, and content-area learning for secondary-school students with learning disabilities, and reading comprehension interventions for students with reading difficulties.

Siuman Chung, MSc, is a Ph.D. student of the Department of Education and Child Studies (specialization Learning Disabilities) at Leiden University, The Netherlands. Her research interests include Curriculum-based Measurement in reading for secondary-school students, and teachers' data-based decision-making.

Nadira Saab, Ph.D., is an Assistant Professor at the ICLON, Leiden University Graduate School of Teaching, The Netherlands. Her research interests involve the impact of powerful and innovative learning methods and approaches on learning processes and learning results, such as collaborative learning, technology enhanced learning, (formative) assessment and motivation.