Manipulating Graph Elements to Assess Preservice Special Educators' Evaluation of Progress Monitoring Data

Cynthia C. Massey,[1] Emily M. Kuntz,[2] Corey Peltier,[2] Mary A. Barczak,[2] and H. Michael Crowson[2]
[1]Georgia Southern University, [2]University of Oklahoma

Abstract

Enhancing special educators' data literacy is critical to informing instructional decision-making, especially for students with learning disabilities. One tool special educators commonly use is curriculum-based measurement (CBM). These data are displayed on time-series graphs, and student responsiveness is evaluated. Graph construction varies and may impact teacher interpretation. This experiment focused on isolating two graphical elements, (a) the presence of an aimline and (b) data points per x- to y-axis ratio (DPPXYR), to determine if they served as analysis-altering elements. Participants, 31 preservice special educators enrolled in two Assessment in Special Education courses, evaluated 48 CBM graphs representing eight data sets with six manipulations. The presence of an aimline significantly increased accuracy in evaluating progress monitoring data, whereas the DPPXYR did not impact decisions. The study outlines the importance of incorporating aimlines into CBM graphs to improve special educators' data literacy, thus enhancing instructional decision-making and the learning outcomes of students with learning disabilities. Further discussion will explore detailed implications for CBM graph construction and use in the classroom.

Keywords: Curriculum-based measurement, effective practice, teacher preparation, data literacy, progress monitoring, graph manipulation, preservice teachers

Data literacy, defined as the ability to interpret and apply data for decision-making (Mandinach & Gummer, 2013), is essential for special educators assessing instructional effectiveness to support students with learning disabilities. In the United States, the Individuals with Disabilities Education Act (IDEA; 2004) states that annual goals must be progress monitored, and, accordingly, special educators evaluate these time series data to determine if the program is designed to ensure the child receives a free and appropriate public education (FAPE; Hott et al., 2020).

This study examined how graph construction may influence preservice special educators' ability to make accurate instructional decisions when evaluating progress monitoring data. The findings have implications for real-life classroom decision-making aimed at improving the lives of students with disabilities, including those with learning disabilities.

Use of Curriculum-Based Measurement for Progress Monitoring

Curriculum-based measurement (CBM) is an evidence-based strategy widely used to support and monitor the academic progress of students with specific learning disabilities. CBMs necessitate that teachers conduct, assess, and interpret the relations between the data and draw informed inferences to make the most appropriate decisions regarding instructional strategies, interventions, and adjustments to meet individual student needs (Espin et al., 2017).

Massey, C. C., Kuntz, E. M., Peltier, C., Barczak, M. A., & Crowson, H. M. (2024). Manipulating Graph Elements to Assess Preservice Special Educators' Evaluation of Progress Monitoring Data. *International Journal for Research in Learning Disabilities, 7*(1), 27. https://doi.org/10.28987/ijrld.7.1.27

Research on CBM may be divided into three stages, each focusing on a different aspect of its application and efficacy (see Fuchs, 2004). Stage 1 is focused on evaluating the psychometric evidence of data collected from a CBM at a single point in time. A typical scenario might involve administering a reading CBM to assess its reliability and how it correlates with the other overall reading achievement measures. Stage 2 is focused on evaluating the psychometric evidence of data collected from a CBM repeatedly across time, thus, an evaluation of the slope. An example would be giving a mathematics CBM weekly for 12 weeks to determine the reliability of performance trends and predict future math achievement. Finally, Stage 3 is focused on instructional utility, for example, evaluating how much training teachers need to administer CBMs with fidelity, evaluate the time-series data to make informed decisions, and then document whether this process yielded stronger student outcomes than when not using CBMs.

This study contributes to the literature pertaining to the third stage, instructional utility, used as a practical tool to improve the educational outcomes of students with learning disabilities. Much of the current literature centers upon psychometric evidence of data at a single time point (i.e., the first stage), while far fewer studies focus on the third stage, instructional utility, and the need for data literacy among educators (Espin et al., 2017; Fuchs et al., 2021; Mandinach & Gummer, 2013; Nelson et al., 2023). Data literacy includes the ability to conduct, assess, and interpret data relationships to make informed decisions to best meet the needs of students with diverse needs, including those with specific learning disabilities (Mandinach & Gummer, 2013). This emphasis on data literacy emphasizes the need for informed decision-making in educational interventions.

Research on Data Literacy and Its Challenges

One framework for using data to inform intervention is data-based individualization (DBI), which focuses on intervention decisions for students who need intensive intervention (National Center on Intensive Intervention [NCII], 2013). Using this model, the instructor identifies a student's area of need, selects an appropriate tool to measure progress, collects baseline data, and establishes student goals for mastery that are both "challenging and ambitious" (*Endrew F. vs. Douglas County School District*, 2017). The educator subsequently provides intervention, collects data on a predetermined frequency schedule (e.g., weekly), and graphs and analyzes the data. Using this process, the instructor makes decisions based on visual data analysis. Additionally, the instructor determines if the intervention or relevant instructional characteristics are related to a change in the student's behavioral performance and estimates the magnitude of this relationship (Filderman et al., 2018; Kratochwill et al., 2013).

While there are various decisions that educators might make based on the CBM progress monitoring data of students with disabilities, this project focused on three foundational preservice teacher decisions to (a) keep the current intervention, (b) decrease intervention intensity, or (c) increase intervention intensity (The IRIS Center, 2015a, 2015b).

Preservice Teachers' Need for Training in Data Literacy

Preservice teachers need time and practice to develop their ability to carry out the DBI process. Research on effective CBM training in teacher preparation programs is limited. In addition, there is a lack of standardized measures to evaluate preservice teachers' skills in this area (Kennedy et al., 2016; Wagner et al., 2017). Without psychometrically sound measures, evaluating preservice educators' knowledge and assessing the effects of training programs is challenging. Also, without adequate instruction, preservice teachers will not be prepared to make effective data-based decisions, impacting their ability to fulfill students' access to an appropriate education.

For this reason, Wagner et al. (2017) recommended sustained training opportunities during teacher preparation to improve preservice teachers' ability to analyze and interpret graphed data visually. To accomplish this, preservice teachers must be taught to visually analyze and interpret a graph as a means to determine the necessary next steps for instruction. This need is recognized in national standards for teaching excellence common in the United States (Council for Exceptional Children's [CEC] Standards for Initial Special Education Preparation, 2015; Collaboration for Effective Educator Development, Accountability, and Reform [CEEDAR] Center's High Leverage Practices, 2014; Council of Chief State School Officers [CCSSO] Interstate Teacher and Support Consortium [InTASC], 2013).

Impact of Graph Construction on Data Analysis

Graph construction is one variable that may facilitate teachers' accuracy in visual analysis. Prior research has reported a large variation in the characteristics of time-series graphs published as part of single-case design studies (Kubina et al., 2021; Ledford et al., 2019; Peltier et al., 2021; Peltier, McKenna, et al., 2022; Peltier, Muharib, et al., 2022) and by practitioners for progress monitoring (Lewis et al., 2021). When interpreting graphed data, these variations can lead to differential decisions. For example, Dart and colleagues (2021) found that time-series graphs produced by four commonly used computer-based progress monitoring systems varied substantially and led to differential estimates of treatment effectiveness. Special educators adjust instruction every day. The lack of standardization in graph construction may impact their decision-making and, ultimately, their students' learning outcomes (Lewis et al., 2021).

On most time-series graphs, the x-axis displays time (e.g., dates, intervention sessions), and the y-axis displays the primary outcome of interest (e.g., academic or behavioral skill). Dart and Radley (2018) suggested a framework for classifying graph elements as either (a) aesthetic-altering or (b) analysis-altering. *Aesthetic-altering elements* change the look of the graph but do not impact the decisions made by a visual analyst. Conversely, *analysis-altering elements* impact the interpretations made by a visual analyst.

There are several aesthetic-altering elements that teachers consider when constructing graphs. First, the color palette, the thickness of lines, and the size and shape of data points all impact the look of the graph. Second, decisions about the axis affect readability, such as using tick marks and setting the interval length between tick marks. Third, labeling the axis, using phase change lines, and providing a key influence on readability is also important. Additionally, data suggest three potential analysis-altering elements to graph construction: (a) y-axis scaling, (b) data points per x-axis to y-axis ratio (DPPXYR), and (c) use of an aimline. The following sections detail each of these elements further.

Y-Axis Scaling

Y-axis scaling involves determining the values and increments of the y-axis, such as the values at which the y-minimum and y-maximum are set. Dart and Radley (2017) demonstrated that the decision to set the y-maximum value impacted visual analysis. Errors in y-axis scaling can lead to potential errors when interpreting the data (Dart & Radley, 2017). Therefore, it is important that special educators protect against y-axis scaling errors when analyzing graphed CBM data. To reduce test threats to the internal validity of our experiment, we opted not to investigate y-axis scaling. Attempting to manipulate this variable and the other two focal variables would have required respondents to evaluate more graphs, possibly leading to testing fatigue.

Data Points Per X- to Y-Axis Ratio (DPPXYR)

The second element that may impact the visual analysis of graphed data is manipulating the DPPXYR (Radley et al., 2018). The DPPXYR extends other recommendations for considering y-axis to x-axis scaling by factoring in data points. The ratio is calculated by dividing the length of the x-axis by the length of the y-axis and then dividing by the number of possible data points plotted along the x-axis. Scholars have suggested that the ratio of the x-axis length compared to the y-axis height may impact visual analysis. Many recommend a time-series graph with an x:y ratio between 8:5 and 3:2 (Cooper et al., 2020). Others suggest considering the density of the data points plotted along the x-axis (Ledford et al., 2019; Radley et al., 2018).

Radley and colleagues (2018) found that the mean DPPXYR for single-case graphs using a multiple-baseline design published in top journals in school psychology was 0.14. They subsequently manipulated graphs to produce a DPPXYR set to 0.14 and +0.5 *SD*, +1.0 *SD*, -0.5 *SD*, and -1.0 *SD.* Type II error rates were higher for the graphs set at +0.5 *SD* and +1.0 *SD,* and Type I error rates were higher for the graphs set at -0.5 *SD* and -1.0 *SD.* These findings led the authors to recommend graphs constructed with DPPXYR between 0.14 and 0.16. The inflated Type I error rates are likely caused by a distorted interpretation of the trend for a truncated x-axis compared to the y-axis height.

Figure 1 provides an example of DPPXYR manipulation. The top graphs (A and B) have a DPPXYR set to 0.05, the middle graphs (C and D) have a DPPXYR set to 0.10, and the bottom graphs (E and F) have a DPPXYR set to 0.15. Readers may have a distorted interpretation of the slope despite the identical data sets, with the top graphs appearing to display a steeper slope than the bottom graphs.

Further replication of DPPXYR manipulations is warranted to determine its impact on visual analysis because, currently, only one experimental study serves as the basis for its recommendation (i.e., Radley et al., 2018). In a recent study, Kuntz et al. (2023) found that the DPPXYR did not significantly impact preservice educators' decision-making accuracy when evaluating progress monitoring graphs. With these mixed results, the current study aimed to expand this work by (a) investigating the impact of DPPXYR on the interpretation of progress monitoring graphs (cf. single-case design graphs) and (b) sam-

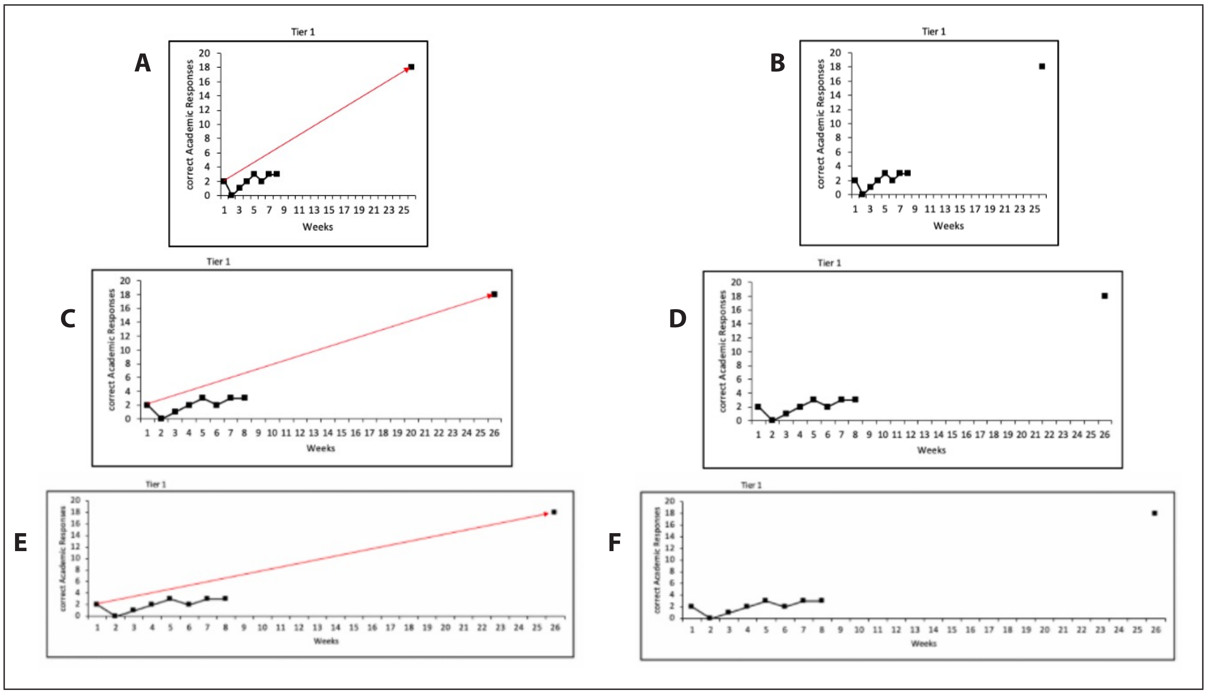

Figure 1
Example Progress Monitoring Graphs With DPPXYR and Aimline Manipulation

Note. (A) DPPXYR = 0.05, with aimline; (B) DPPXYR 0.05, no aimline; (C) DPPXYR = 0.10, with aim line; (D) DPPXYR = 0.10, no aimline; (E) DPPXYR = 0.15, with aimline; (F) DPPXYR = 0.15, no aimline.

pling a different population of students – preservice special educators with additional training in evaluating progress monitoring graphs (cf. all preservice educators with limited training).

Aimline

The third potential analysis-altering element is the use of an aimline, a characteristic unique to time-series graphs for progress monitoring. The aimline is typically created by extending a straight line connecting a baseline datum to a set criterion at a specified time point in the future. Adding the graphical elements of an aimline provides visual support to estimate the linear progress a student must make to obtain the end-of-year goal. Figure 1 provides an example of aimline manipulation for each of the aforementioned DPPXYR manipulations.

Prior work has not been able to parse out the impact of an aimline because multiple other graphical characteristics varied across the computer-based progress monitoring software used (i.e., Dart et al., 2021). Kuntz et al. (2023) identified a 20% increase in accuracy of decision-making for novice preservice educators. We investigated whether these findings would replicate with preservice special educators who had more training in progress monitoring visual analysis. Such an inquiry is essential because special educators who are better prepared to interpret student data are more likely to facilitate improved student outcomes.

Purpose of the Current Study

The study extended previous work by investigating if the DPPXYR and aimline impacted accurate decisions made by preservice educators. The following research questions guided the study:

1. What is the accuracy rate of preservice special educators' intervention decision-making when evaluating progress monitoring data? Does this rate vary based on the data set?
2. Does the presence of an aimline produce more accurate intervention decisions among preservice special educators?
3. Does the manipulation of the DPPXYR impact the accuracy of intervention decision-making among preservice special educators?
4. Is there an interaction between the presence of an aimline and the DPPXYR in the accurate intervention decision-making among preservice special educators?

Method

Participants and Courses

The study sample included preservice special educators enrolled in an assessment course required for a degree program. Recruitment occurred in the fall of 2020 at two universities in the southern United States. Of the 37 students enrolled, 31 consented and provided complete data sets. The mean age was 21.3 (*SD* = 2.44; range = 19 to 33). Most participants identified as Women (*n* = 29, 94%) and White (*n* = 28, 90%; Black: *n* = 1; Latino: *n* = 1; American Indian or Alaska Native: *n* = 1). Data were distributed between two university sites, with 14 participants at Site 1 (*n* = 14) and 17 at Site 2 (*n* = 17).

The assessment course at the two institutions shared a similar focus, but participants' prior learning histories differed at the two sites. Each course provided an understanding of the legislation, policies, and procedures pertaining to the eligibility and assessment of students with disabilities. Within these topics, students were taught how to administer and analyze various evidence-based assessment tools. Both course curricula covered a range of topics, including evaluation techniques outlined in IDEA and strategies for providing comprehensive education with an emphasis on data-based decision-making and progress monitoring to improve instructional outcomes.

Course instructors included instruction on best practices in CBM, progress monitoring, error analysis, data collection, the process for graphing data, and conducting effective visual analysis (Hosp et al., 2016). As part of the courses, the instructors utilized the case studies modules from the IRIS Center (The IRIS Center, 2015a, 2015b), which offered in-depth insights into this topic, encompassing examples of graphs and data sets for hands-on exercises. Following the in-class instruction, hands-on practice, and out-of-class independent practice assignments, students completed a survey upon which this study is based. It served as an exercise to assess their ability to make accurate instructional decisions based on the learned concepts.

At University Site 1, participants were in the fall semester of their senior year. They were concurrently completing a semester-long practicum experience (i.e., their second field placement consisting of approximately 170 hours) in a secondary setting (either middle or high school) paired with a special education cooperating teacher certified in mild/moderate disabilities, including learning disabilities. During the semester, participants completed a project that required them to identify an appropriate CBM for reading and mathematics and administer these measures to selected students weekly for eight weeks. They created appropriate aimlines, scored, graphed, and interpreted the progress monitoring data. Prior to this study, in the previous spring semester, the participants had completed a semester-long practicum experience (i.e., their first field placement consisting of approximately 170 hours) in an elementary special education setting. In these elementary settings, they gained experience administering CBMs related to early literacy and reading and providing reading intervention to students.

At University Site 2, the participants were in the first semester of their junior year and were completing a 30-hour practicum experience in an elementary school placement with a special education clinical supervisor. As part of the assessment course, participants were introduced to CBMs through multi-week, face-to-face sessions encompassing reading, mathematics, spelling, and writing. Instruction included participants studying case studies, analyzing single-case data, creating graphs, plotting data, developing aimlines, and scoring, graphing, and interpreting simulated progress monitoring data.

Procedures

Following Institutional Review Board approval, a recruitment message and consent form were included in the online survey, which students completed as an assignment after the CBM unit of study. Course instructors were blinded from knowing who provided consent as another researcher not affiliated with the courses took responsibility for recruitment to prevent coercion. When completing the course survey assignment, students who provided consent were routed to a page that collected demographic information before proceeding to the survey. Students who opted out of the study were routed directly to the course assignment survey. A researcher unaffiliated with the courses collected all responses, de-identified consenting participants' data, and reported assignment completion to the students' corresponding instructor for grading.

Survey Instrument

Participants evaluated 48 graphs. Each survey page included one graph above a response option. The item was, "Given the student's current performance, what instructional decision do you feel is needed?" Response options included *keep intervention intensity, increase intervention intensity,* and *decrease intervention intensity* (The IRIS Center, 2015a, 2015b; Kuntz et al., 2023; NCII, 2014).

Data Sets

The research team developed eight data sets reflecting progress monitoring data. To control the task, all data sets contained the same dependent variable – correct academic responses presented as count data. The data sets aligned with the recommended administration and scoring protocol for vocabulary CBMs (see Hosp et al., 2016). We selected the variable "correct academic responses" with a maximum of 20 because of its generalizability across age bands, content areas, and student populations.

The eight data sets were altered based on the following characteristics. Four data sets presented data with two different phases of instruction, labeled Tier 1 and Tier 2. Each phase included eight data points (i.e., 16 total data points). The remaining four data sets depicted data with one instructional phase, labeled either Tier 1 or Tier 2. Each of these data sets included eight data points in total. Regarding correct instructional decisions (i.e., correct participant responses), four data sets clearly depicted a data pattern with *keep intervention intensity* as the correct response, three data sets depicted a data pattern with *increase intervention intensity* as the correct response, and one data set depicted a data pattern with *decrease intervention intensity* as the correct response.

Several approaches to evaluating progress monitoring data are commonly used. In this study, students were instructed to adhere to the learned Steps in the DBI Process framework developed by the National Center on Intensive Intervention provided within The IRIS Center's modules on Data-Based Individualization, Part 1 and Part 2 (The IRIS Center, 2015a, 2015b; Kuntz et al., 2023; NCII, 2014). All data sets were developed so there was no ambiguity about which instructional decision was correct.

The following responses indicated the correct decision: If the last three data points in the data set were below the aimline, the correct response was coded as *increase intervention intensity*. If the last three data points were at or slightly above the aimline, the correct decision was coded as *keep intervention intensity*. If the last three data points were below the aimline, the correct decision was coded as *decrease intervention intensity* (The IRIS Center, 2015a, 2015b; Kuntz et al., 2023; NCII, 2014).

Graphs

All graphs depicted the x-axis as time and the y-axis as correct academic responses. The x-axes were labeled as weeks, scaled from 1 to 26, and every odd value was labeled (e.g., 1, 3, 5). The y-axes were labeled as correct academic responses, scaled from 0 to 20, and every even value was labeled (e.g., 0, 2, 4). The x- and y-axes were black with a 1pt line thickness. Lines connecting data points were black with a 1.5pt thickness. A solid, vertical phase change line indicated where the instruction change occurred for data with two instructional phases. The x- and y-axes included tick marks presented outside the graph space. Each datum was displayed in black, representing a solid square in 6pt font.

For each data set, the two variables manipulated were the DPPXYR and the aimline. First, there were three manipulations of the DPPXYR (i.e., .05, .10, .15); second, each DPPXYR value was graphed with and without an aimline (i.e., present, not present). In total, we created six graphs for each of the eight data sets ($N = 48$). When the aimline was present, we displayed it in red with a 0.5pt line thickness (see Figure 1). The aimline was created by connecting the first datum to the end-of-year goal (i.e., 18 out of 20 correct academic responses, which is a 90% accuracy criterion).

Data Analysis

The first research question examined the accuracy of preservice special educators' decisions when evaluating progress monitoring graphs. Descriptive data were calculated at both the graph and participant level. Research Questions 2, 3, and 4 investigated whether the presence of an aimline and the DPPXYR independently and/or interactively would predict the probability of a participant making a correct decision.

The aimline variable was coded according to whether an aimline was present or absent in a particular graph (i.e., 0 = absent, 1 = present). The DPPXYR variable was originally coded as 0.05, 0.10, and 0.15. It was re-coded into two dummy variables for the regression analyses, with 0.05 serving as the reference category. Based on this coding, the regression slopes associated with each dummy variable are interpreted as a difference in predicted log odds between the 0.10 or 0.15 graph manipulations and the reference category of 0.05. Similarly, the regression slope associated with the aimline variable can be interpreted as the predicted difference in log odds when exposed to a graph containing an aimline and one not containing an aimline. The questions were presented in a mixed order to prevent participants from discerning any patterns in the answers, thereby eliminating the possibility of order effects. Participants' decisions were coded as 0 = incorrect or 1 = correct.

Given that participants responded to 48 separate graphs, an analytic strategy was required that accounted for both the binary outcome and the repeated measurement within participants. Based on those considerations, we tested the independent and interactive effects of aimline and DPPXYR on the probability of making a correct judgment for a graph using multilevel binary logistic regression.

We began our analysis by examining an unconditional model to evaluate the presence of clustering in our data. Next, we tested two substantive models involving our predictors. First, we regressed the binary outcome variable for the first model onto the aimline and DPPXYR dummy variables. Then, we re-specified the model to include interaction terms to test for an interaction effect of our within-subject factors on the probability of making a correct judgment. Next, we compared the fit of our two substantive models to arrive at a final preferred model. After deciding upon our preferred model, we re-specified it to test for between-university differences in its effects. All multilevel analyses were performed using the 'lme4' (Bates et al., 2015), 'lmerTest' (Kuznetsova et al., 2017), and 'jtools' (Long, 2022) packages in Rstudio.

Results

Accuracy of Decision-Making Across Data Sets

Across all graphs, the mean correct response was 69% (*SD* = 13%). Participants from University Site 1 averaged 71% correct responses (*SD* = 11%); participants from University Site 2 averaged 68% correct responses (*SD* = 15%). These differences were not statistically significant. We also reviewed correct responses across our graph-altering variables (see Table 1). For graphs with the presence of an aimline, participants responded correctly 73% (*SD* = 14%) of the time, and for graphs without the presence of an aimline, they responded correctly 66% (*SD* = 15%) of the time. For graphs with a DPPXYR of 0.15 and an aimline, participants responded correctly 68% (*SD* = 17%) of the time. For graphs with a DPPXYR of 0.15 and no aimline, they responded correctly 65% (*SD* = 17%) of the time. For graphs with a DPPXYR of 0.10 and an aimline, participants responded correctly 75% (*SD* = 17%) of the time. For graphs with a DPPXYR of 0.10 and no aimline, they responded correctly 65% (*SD* = 15%) of the time. Finally, for graphs with a DPPXYR of 0.05 and an aimline, participants responded correctly 75% (*SD* = 16%) of the time. For graphs with a DPPXYR of 0.05 and no aimline, they responded correctly 67% (*SD* = 19%) of the time.

There was large variability in the accuracy rate at the participant level (range = 38% to 83%). When analyzing the distribution, three participants were accurate 80% of the time or more, 17 participants were accurate between 70% and 79%, five participants were accurate between 60% and 69%, one participant was accurate between 50% and 59%, four participants were accurate between 40% to 49%, and one participant was accurate below 40% of the times. Variability in the accuracy of responses differed by correct instructional decision. The average accuracy rates were the least variable and greatest for the graphs containing data sets with a correct response of *keep intervention intensity* – 66%, 75%, and 88%. For the graphs containing data sets with a correct response of *increased intervention intensity,* the accuracy rates were more variable – 45%, 70%, 83%, and 91%. The data set with a correct response of *decrease intervention intensity* had a mean accuracy of 40%.

Impact of Graph Manipulations on Accuracy

As noted earlier, we began our analysis by evaluating a random-intercept model containing no predictors. The intraclass correlation coefficient (ICC) for the model – computed as ICC = (var(μ_0))/(var(μ_0)+(π^2/3) (see Sommet & Morselli, 2017) – was 0.0813, providing evidence of non-trivial clustering in our data. We compared the fit of this model to that of a standard (single level) logistic regression using a likelihood ratio chi-square test. The result of this test, LR $\chi(1) = 47.89$, $p < 0.001$, indicated that our model specifying randomly varying intercepts fit the data significantly better than a single-level model. This finding, in conjunction with the ICC, indicated substantial between-participant variation in the probability (expressed in the metric of log odds) of correct responses to the graphs. Given the evidence of clustering, we moved on to testing our substantive models of interest. Results from those models are provided in Table 2.

For the first substantive model (i.e., Model 1), we regressed the binary dependent measure (1 = correct, 0 = incorrect) onto the aimline and DPPXYR variables. A likelihood ratio test revealed that this model fit the data significantly better than our previous intercept-only model, LR $\chi(3) = 19.854$, $p < 0.001$, aimline emerged as a positive and significant predictor ($b = .5267$, $SE = .1193$, $p < 0.001$) of the log odds of a participant making a correct judgment from a graph that was presented to them. The odds ratio for aimline was 1.6933, indicating that when a graph was shown containing an aimline, the odds of a correct judgment was (100%) [1.6933-1] = 69.33% greater than the odds

Table 1
Descriptive Results of Participant Responses for Each of the 48 Graphs by University Site

			University site 1 (*n* = 14)			University site 2 (*n* = 17)		
Data set	Aimline	DPPXYR	Keep	Increase	Decrease	Keep	Increase	Decrease
3	0	0.05	57.1%	**35.7%**	7.1%	70.6%	**29.4%**	0.0%
2	0	0.05	**78.6%**	7.1%	14.3%	**82.4%**	11.8%	5.9%
4	0	0.05	21.4%	**78.6%**	14.3%	17.6%	**76.5%**	5.9%
5	0	0.05	**85.7%**	7.1%	7.1%	**82.4%**	5.9%	11.8%
6	0	0.05	0.0%	**100.0%**	0.0%	17.6%	**76.5%**	5.9%
7	0	0.05	71.4%	0.0%	**28.6%**	41.2%	5.9%	**52.9%**
1	0	0.05	**71.4%**	0.0%	28.6%	**35.3%**	17.6%	47.1%
8	0	0.05	7.1%	**92.9%**	0.0%	5.9%	**82.4%**	11.8%
1	0	0.10	**71.4%**	0.0%	28.6%	**41.2%**	11.8%	47.1%
2	0	0.10	**78.6%**	7.1%	14.3%	**88.2%**	11.8%	0.0%
5	0	0.10	**64.3%**	21.4%	14.3%	**76.5%**	11.8%	11.8%
8	0	0.10	7.1%	**92.9%**	0.0%	0.0%	**94.1%**	5.9%
3	0	0.10	57.1%	**35.7%**	7.1%	70.6%	**23.5%**	5.9%
4	0	0.10	14.3%	**85.7%**	0.0%	23.5%	**70.6%**	5.9%
7	0	0.10	71.4%	7.1%	**21.4%**	82.4%	0.0%	**17.6%**
6	0	0.10	7.1%	**92.9%**	0.0%	0.0%	**94.1%**	5.9%
7	0	0.15	71.4%	7.1%	**21.4%**	88.2%	5.9%	**5.9%**
6	0	0.15	7.1%	**92.9%**	0.0%	17.6%	**76.5%**	5.9%
8	0	0.15	7.1%	**92.9%**	0.0%	5.9%	**88.2%**	5.9%
1	0	0.15	**71.4%**	7.1%	21.4%	**70.6%**	11.8%	17.6%
3	0	0.15	57.1%	**35.7%**	7.1%	64.7%	**29.4%**	5.9%
4	0	0.15	21.4%	**71.4%**	7.1%	17.6%	**76.5%**	5.9%
5	0	0.15	**57.1%**	21.4%	21.4%	**76.5%**	11.8%	11.8%
2	0	0.15	**85.7%**	7.1%	7.1%	**88.2%**	11.8%	0.0%
8	1	0.05	7.1%	**92.9%**	0.0%	11.8%	**88.2%**	0.0%
2	1	0.05	**92.9%**	7.1%	0.0%	**94.1%**	0.0%	5.9%
6	1	0.05	21.4%	**78.6%**	0.0%	35.3%	**58.8%**	5.9%
1	1	0.05	**64.3%**	0.0%	35.7%	**70.6%**	0.0%	29.4%
4	1	0.05	7.1%	**92.9%**	0.0%	5.9%	**88.2%**	5.9%
7	1	0.05	57.1%	0.0%	**42.9%**	11.8%	11.8%	**76.5%**
5	1	0.05	**78.6%**	0.0%	21.4%	**70.6%**	5.9%	23.5%
3	1	0.05	28.6%	**71.4%**	0.0%	41.2%	**47.1%**	11.8%
5	1	0.10	**64.3%**	7.1%	28.6%	**82.4%**	0.0%	17.6%
7	1	0.10	50.0%	0.0%	**50.0%**	29.4%	5.9%	**64.7%**
4	1	0.10	0.0%	**100.0%**	0.0%	11.8%	**82.4%**	5.9%
3	1	0.10	50.0%	**50.0%**	0.0%	52.9%	**47.1%**	0.0%
1	1	0.10	**78.6%**	0.0%	21.4%	**64.7%**	17.6%	17.6%
8	1	0.10	0.0%	**100.0%**	0.0%	5.9%	**88.2%**	5.9%
6	1	0.10	7.1%	**92.9%**	0.0%	5.9%	**88.2%**	5.9%
2	1	0.10	**78.6%**	21.4%	0.0%	**76.5%**	17.6%	5.9%
6	1	0.15	14.3%	**85.7%**	0.0%	23.5%	**64.7%**	11.8%
5	1	0.15	**78.6%**	0.0%	21.4%	**82.4%**	0.0%	17.6%
8	1	0.15	0.0%	**100.0%**	0.0%	0.0%	**88.2%**	11.89%
7	1	0.15	57.1%	0.0%	**42.9%**	41.2%	0.0%	**58.8%**
3	1	0.15	14.3%	**78.6%**	7.1%	35.3%	**64.7%**	0.0%
2	1	0.15	**78.6%**	21.4%	0.0%	**88.2%**	5.9%	5.9%
1	1	0.15	**71.4%**	0.0%	28.6%	**82.4%**	5.9%	11.8%
4	1	0.15	7.1%	**92.9%**	0.0%	11.8%	**82.4%**	5.9%

Note. Aimline codes are 0 = No aimline and 1 = aimline; DPPXYR = Data points per x-y axis ratio. **Correct responses are in bold**.

Table 2
Mixed Model Results

	Model 0 (null) B / Exp(B)	Model 1 B / Exp(B)	Model 2 B / Exp(B)
Level 1 predictors			
Intercept	.9521*** / 2.591	.7320***/2.0793	.7882*** / 2.1993
Aimline		.5267*** / 1.6933	.4034 *** / 1.4969
.10 (Dum 1)		-.0364 / .9386	-.1163 / .8902
.15 (Dum 2)		-.0213 / 9789	-.1354 / .8734
.10 (Dum 1) X aimline			.1164 / 1.1234
.15 (Dum 2) X aimline			2541 / 1.2893
Variance components			
var()	.2900	.3000	.3003
Marginal R-square	.0000	.0192	.0199
Conditional R-square	.0810	.1011	.1019
ICC	.0810	.0836	.0837
AIC	1747.1815	1733.3279	1736.5769
BIC	1757.7919	1759.8538	1773.7132
Δdf		3	2
LR χ^2		19.854***	.6191

Note. Dum = Dummy-coded variable, var = Variance, ICC= Intraclass correlation coefficient, AIC = Akaike information criterion, BIC = Bayesian information criterion, Δdf = Change in degrees of freedom, LR χ^2 = Likelihood ratio chi-square. B / Exp(B) refers to the unstandardized regression slope and exponentiated regression slope, respectively. Exponentiated regression slopes are interpreted as odds ratios. ***p<.001. LR χ^2 tests the difference in fit between adjacent models [Model 0 vs. Model 1; Model 1 vs. Model 2].

of correct judgment when a graph was shown without an aimline. Neither regression slope associated with the dummy variables for DPPXYR was statistically significant. The marginal and conditional R-squares – computed using the method of Nakagawa and Schielzeth (2013) – were 0.0192 and 0.1011, respectively (where the former is the variance accounted for by the fixed effects and the latter is accounted for by the fixed and random effects).

Next, we re-specified the model (i.e., Model 2) to include the previous independent variables and their interaction. Two interaction terms were formed as the product of each DPPXYR dummy variable and the aimline variable. As before, our model fit the data significantly better than the random-intercept-only model, LR $\chi(5) = 20.605$, $p < 0.001$. Nevertheless, it did not fit significantly better, LR $\chi(2) = 0.7509$, $p = 0.687$, than Model 1, containing only the aimline and DPPXYR dummy variables. The only (near) significant predictor in Model 2 containing the interaction terms was aimline ($b = 0.4034$, $SE = 0.2064$, $p = 0.0507$). The marginal and conditional R-squares for Model 2 were 0.0199 and 0.1019, respectively. Given that Model 2 failed to significantly improve upon the fit of our model, we retained Model 1 as our preferred model moving forward.

Although tertiary to our main analyses, we re-specified Model 1 to test for the possibility that the effects in our retained model might vary between participants attending the two universities. First, we tested a model that included university as a Level 2 predictor of variation in participants' random intercepts. A likelihood ratio test revealed that this model did not fit the data significantly better than our original Model 1, LR $\chi(1) = 0.439$, $p = 0.5076$. Next, we tested a model that included cross-level interaction terms to test whether the slopes for the aimline and DPPXYR dummy variables might vary between students attending the two universities. Due to model convergence problems by allowing slopes for the within-subjects effects to vary randomly, our final parameterization did not involve estimating variance components for random slopes. The likelihood ratio test comparing this interaction model against our original Model 1 was not statistically significant, LR $\chi(4) = 0.7898$, $p = 0.9398$.

To provide a more detailed look at the effects of the graph manipulations, we computed the probability of a participant making a correct judgment depending on the DPPXYR condition and the presence or absence of an aimline. Table 3 contains the expected probability of a correct judgment as a function of the two within-person manipulations. These probabilities were computed from the fixed effect portion of Model 1, given as ln(Y = 1) = 0.7320 + 0.5267*aimline - 0.0364*Dum1 - 0.0213*Dum2 (i.e., Dum1 = 0.10, Dum2 = 0.15). Based on the Model 1 estimates, the expected probability of a person making a correct judgment on a graph ranged from 0.661 to 0.779. As one would expect based on the results described earlier, the probability of making a correct judgment was greater when the aimline was present in a graph than when it was absent. Averaged across DPPXYR conditions, the probability of making a correct decision when an aimline was present was approximately 1.157 times greater than when the aimline was absent.

Discussion

Data literacy is essential for special educators as it directly impacts the success of students with learning disabilities and those who learn differently. Legal mandates in the United States require educators to monitor progress toward students' Individualized Education Program (IEP) goals and ensure students make adequate progress toward those goals (*Endrew F. vs. Douglas County School District,* 2017; IDEA, 2004). CBMs are an effective and practical way to measure that progress, yet research has identified a gap in preservice special educators' ability to interpret graphed data accurately (Kennedy et al., 2016; Lane et al., 2021; van den Bosch et al., 2017; Wagner et al., 2017).

This study extends emerging literature by examining the effect of the DPPXYR and an aimline as analysis-altering elements. Specifically, we evaluated how manipulating (a) an aimline and (b) DPPXYR impacted the accuracy of preservice special educators' visual analysis. The findings rendered outcomes that can be used to direct future research and practice, specifically in special education, where academic and behavior monitoring is essential to determine students' progress toward goal mastery.

Overall, accurate decision-making was 69%, with considerable variability across preservice special educators (range = 38% to 83%). While better than chance (i.e., 33%), this accuracy rate is somewhat concerning and indicates that while initial training is beneficial, more training and practice is needed for educators to reach a level of mastery. Moreover, the range underscores the need to investigate training models to identify critical components that enhance accurate evaluation, and given the variability across educators, training models must differentiate training intensity.

One possible component may be a tiered support system (similar to response to intervention [RtI] procedures) targeting educators demonstrating low accuracy with differentiated instruction. An interesting and exploratory finding showcased differential accurate rates based on the data patterns (e.g., patterns indicating a need to keep intervention intensity) presented to educators. Perhaps explicitness on when to decrease intervention intensity versus maintain intervention intensity would support correct response decisions.

Pertaining to graphical elements that could improve accuracy, an aimline was a statistically significant predictor of preservice special educators' correct responses. First, the odds ratio indicated that the presence of an aimline led to 69.33% greater accuracy than a graph without an aimline. This finding was more than three times larger for more novice preservice educators (see Kuntz et al., 2023). Second, manipulating the DPPXYR, with or without an aimline, did not prove to be statistically significant. This finding also corresponds with our previous findings (Kuntz et al., 2023). Combined, these two outcomes indicate that the aimline influenced preservice special educators more than the DPPXYR and support the notion that an aimline is a potential analysis-altering element. Thus, the findings from our research underline the need for teacher training programs to emphasize the use and interpretation of aimlines. Such instruction will better equip future educators to evaluate student progress and make more informed decisions on intervention strategies, including those used during RtI procedures, ultimately enhancing learning outcomes for students with learning disabilities.

Table 3
Predicted Probability of Correct Judgment

DPPXYR	Aimline not present	Aimline present
.05	.675	.779
.10	.661	.768
.15	.671	.775

Note. The main values are computed based on Model 1 fixed-effects parameter estimates.

Limitations and Future Research

Findings should be interpreted in lieu of the following limitations: (a) an unrepresentative and small sample of participants, (b) forced choice on the decision options from which participants could select, and (c) selected data sets may not represent the most frequently collected data sets in practice.

First, the sample was recruited from only two cohorts of preservice special educators at two separate institutions. The sample was small due to small cohorts of special educators at each institution and may not represent the target population (i.e., preservice special educators). Though not statistically significant, prior learning histories with CBMs, the graphing of time-series data, and the visual analysis of progress monitoring graphs may have some differential influence on responses across institutions based on curriculum design and when faculty prepare preservice educators to evaluate student knowledge in their program of study. Future research may aim to control for prior learning experiences to determine how this informs decision-making accuracy and whether visual analysis training may need to differ based on the experience level of preservice educators.

Second, future researchers may consider whether treating the dependent variable dichotomously (i.e., correct or incorrect) is the best approach. This decision impacted the sensitivity of our analysis. However, There are not values in parentheses. was made because, ultimately, teachers will likely face a forced-choice decision in practice to *keep* intervention, *increase* intervention intensity, or *decrease* intervention intensity (The IRIS Center, 2015a, 2015b; Kuntz et al., 2023; NCII, 2014). Future work can aim to evaluate the categorical decision (i.e., keep, increase, decrease), use a more sensitive measure (e.g., the magnitude of student response), and perhaps collect think-aloud data to triangulate the process for making decisions (see Espin et al., 2017). Revised measures may allow researchers to interpret differential decisions from preservice educators more purposefully.

Third, the participants evaluated simulated data sets that may not represent data collected in practice. Furthermore, these data were presented without anecdotal data and contextual information commonly present in practice (e.g., attendance rates, observation during the testing session, behavioral difficulties). Anecdotal and contextual influence may impact teacher decision-making, which could increase or decrease accuracy in data-based decision-making. While these pseudo situations provide evidence of the accuracy rate of decisions and graph construction however, future work on these decisions using actual data sets familiar to educators is needed. Fortunately, other researchers are beginning to investigate the social validity of simulated data sets in addition to evaluating the effectiveness of training visual analysis of time-series graphs (see Lane et al., 2021). Future investigations could combine the results of this work and the use of graphing manipulations to improve visual analysis.

Last, although the data patterns followed recommended administration and scoring protocols for vocabulary CBM (see Hosp et al., 2016), other experts beyond the research team did not vet the data patterns or resulting graphs. Expert confirmation of the correct intervention decision would enhance the validity of the scores collected from our instrument. Future research could assess the validity of the data with an expert panel and perhaps collect data patterns confirmed to align with practice (Lane et al., 2021). Furthermore, the aimline was set linearly (i.e., expecting linear growth). As researchers begin investigating other models to evaluate the rate of improvement (i.e., parametric, growth curve modeling), it is worth investigating if this impacts the benefits of the aimline.

Implications for Practice

The results from this study yield several implications for improving the education of students with learning disabilities. First, the study adds to the literature regarding preservice special educators' accuracy in evaluating CBM data. Accuracy rates varied following a unit on CBM and highlighted the need for more intensive training for some preservice teachers. This suggests that an initial training module, while beneficial, may not suffice for mastery, and that ongoing training is needed to improve proficiency. Accuracy improved for graphs with aimlines, demonstrating their potential benefit in determining the next steps for instruction. Thus, preservice teachers should be taught to construct graphs using aimlines. Second, some specific data patterns yielded less accurate decisions. Specifically, participants were least accurate when data indicated growth beyond the aimline (i.e., a case to decrease intervention intensity), and the greatest variability occurred when data indicated insufficient growth (i.e., a case to increase intervention intensity). Faculty preparing preservice teachers should ensure varied data sets are used during training to enhance the generalization of skills and potential confidence in changing instruction for students.

Conclusion

This research indicates the important role aimlines can play in improving the accuracy of preservice special educators' assessment of progress monitoring data. This discovery further proves that this graphical element should be categorized as analysis-altering. In this study, the DPPXYR did not significantly influence decision-making. Specific data patterns may impact accuracy, such as when to maintain instruction. However, the variability when determining to decrease and increase the intervention intensity demonstrates the need for further training on graph analysis within special education educator preparation programs. This research has specific implications for improving outcomes for students with learning disabilities and difficulties. The enhanced training is especially relevant for teacher preparation programs as it could facilitate more accurate CBM decisions, thereby creating more effective learning environments for all students with disabilities.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Collaboration for Effective Educator Development, Accountability, and Reform [CEEDAR]. (2014). *High leverage practices*. https://ceedar.education.ufl.edu/high-leverage-practices/

Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis* (3rd ed.). Pearson.

Council for Exceptional Children [CEC]. (2015). *Initial special education preparation standards (Standard No. 4.2)*. https://exceptionalchildren.org/standards/initial-special-education-preparation-standards

Council of Chief State School Officers [CCSSO]. (2013). *InTASC model core teaching standards and learning progressions for teachers 1.0.* (Standard No. 6c). https://ccsso.org/sites/default/files/20172/2013_INTASC_Learning_Progressions_for_Teachers.pdf

Dart, E. H., & Radley, K. C. (2017). The impact of ordinate scaling on the visual analysis of single-case data. *Journal of School Psychology, 63*, 105–118. https://doi.org/10.1016/j.jsp.2017.03.008

Dart, E. H., & Radley, K. C. (2018). Towards a standard assembly of linear graphs. *School Psychology Quarterly, 33*(3), 350–355. https://psycnet.apa.org/doi/10.1037/spq0000269

Dart, E. H., Van Norman, E. R., Klingbeil, D. A., & Radley, K. C. (2021). Graph construction and visual analysis: A comparison of curriculum-based measurement vendors. *Journal of Behavioral Education*. https://doi.org/10.1007/s10864-021-09440-7

Endrew F. versus Douglas County School District. (2017). 137 S.Ct. 988, 1001. Supreme Court of the United States. https://www.supremecourt.gov/opinions/16pdf/15-827_0pm1.pdf

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. https://doi.org/10.1111/ldrp.12123

Filderman, M. J., Toste, J. R., Didion, L. A., Peng P., & Clemens, N. H. (2018). Data-based decision making in reading interventions: A synthesis and meta-analysis of the effects for struggling readers. *The Journal of Special Education, 52*(3), 174–187. https://doi.org/10.1177/0022466918790001

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*(2), 188–192. https://doi.org/10.1080/02796015.2004.12086241

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (2021). Bringing data-based individualization to scale: A call for the next-generation technology of teacher supports. *Journal of Learning Disabilities, 54*(5), 319–333. https://doi.org/10.1177/0022219420950654

Hott, B. L., Morano, S., Peltier, C., Pulos, J., & Peltier, T. (2020). Are students with mathematics learning disabilities receiving FAPE? Insights from a descriptive review of individualized education programs. *Learning Disabilities Research & Practice, 35*(4), 170–179. https://doi.org/10.1111/ldrp.12231

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2016). How to conduct content-area CBM. In M. K. Hosp, J. L. Hosp, & K. W. Howell (Eds.), *The ABCs of CBM: A practical guide to curriculum-based measurement* (2nd ed., pp. 167–174). Guilford Publications.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400, (2004).

The IRIS Center. (2015a). *Intensive intervention (Part 1): Using data-based individualization to intensify instruction*. Retrieved from https://iris.peabody.vanderbilt.edu/module/dbi1/

The IRIS Center. (2015b). *Intensive intervention (Part 2): Collecting and analyzing data for data-based individualization*. Retrieved from https://iris.peabody.vanderbilt.edu/module/dbi2/

Kennedy, M. J., Wagner, D., Stegall, J., Lembke, E., Miciak, J., Alves, K. D., Brown, T., Driver, M. K., & Hirsch, S. E. (2016). Using content acquisition podcasts to improve teacher candidate knowledge of curriculum-based measurement. *Exceptional Children, 82*(3), 303–320. http://doi.org/10/1177/0014402915615885

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*(1), 26–38. https://doi.org/10.1177/0741932512452794

Kubina, R. M., Kostewicz, D. E., King, S. A., Brennan, K. M., Wertalik, J., Rizzo, K., & Markelz, A. (2021). Standards of graph construction in special education research: A review of their use and relevance. *Education and Treatment of Children, 44,* 275–290. https://doi.org/10.1007/s43494-021-00053-3

Kuntz, E., Massey, C. C., Peltier, C. Barczak, M., & Crowson, H. M. (2023). Graph manipulation and the impact on pre-service teachers' accuracy in evaluating progress monitoring data. *Teacher Education and Special Education, 46*(1), 65–82.

Kuznetsova A., Brockhoff, P. B., & Christensen, R.H.B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lane, J. D., Shepley, C., & Spriggs, A. D. (2021). Issues and improvements in the visual analysis of A-B single case graphs by pre-service professionals. *Remedial and Special Education, 42*(4), 235–247. https://doi.org/10.1177/0741932519873120.

Ledford, J. R., Barton, E. E., Severini, K. E., Zimmerman, K. N., & Pokorski, E. A. (2019). Visual display of graphic data in single case design studies: Systematic review and expert preference analysis. *Education and Training in Autism and Developmental Disabilities, 54*(4), 315–327. https://osf.io/sqm7r/download

Lewis, H.H.J., Radley, K. C., & Dart, E. H. (2021). The effect of graph standardization on intervention evaluation of practitioner-created graphs. *Psychology in the Schools.* Advanced online publication. https://doi.org/10/1002/pits.22621

Long, J. A. (2022). *jtools: Analysis and presentation of social scientific data.* R package, version 2.2.0. https://cran.r-project.org/package=jtools.

Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator prep. *Educational Researcher, 42*(1), 30–37. https://doi.org/10.3102/0013189X12459803

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

National Center on Intensive Interventions (NCII). (2013). *Data-based individualization: A framework for intensive intervention.* U.S. Department of Education, Office of Special Education. Retrieved from http://www.intensiveintervention.org/sites/default/files/DBI_Framework.pdf

National Center on Intensive Intervention (NCII). (2014). *Informal academic diagnostic assessment: Using data to guide intensive instruction: Reviewing graphed data.* U.S. Department of Education, Office of Special Education Programs. Retrieved from https://intensiveintervention.org/resource/informal-academic-diagnostic-assessment-using-data-guide-intensive-instruction-part-2

Nelson, G., Kiss, A. J., Codding, R. S., McKevett, N. M., Schmitt, J. F., Park, S., ... & Hwang, J. (2023). Review of curriculum-based measurement in mathematics: An update and extension of the literature. *Journal of School Psychology, 97*, 1–42. https://doi.org/10.1016/j.jsp.2022.12.001

Peltier, C., McKenna, J. W., Sinclair, T. E., Garwood, J., & Vannest, K. J. (2022). Brief report: Ordinate scaling and axis proportions of single-case graphs in two prominent EBD journals from 2010–2019. *Behavioral Disorders, 47*(2), 134–148. https://doi.org/10.1177/0198742920982587

Peltier, C., Morano, S., Shin, M., Stevenson, N., & McKenna, J.W. (2021). A decade review of single-case graph construction in the field of learning disabilities. *Learning Disability Research and Practice, 36*(2), 121–135. https://doi.org/10.1111/ldrp.12245

Peltier, C., Muharib, R., Haas, A., & Dowdy, A. (2022). A decade review of two potential analysis altering variables in graph construction. *Journal of Autism and Developmental Disorders, 52*, 714–724. https://doi.org/10.1007/s10803-021-04959-0

Radley, K. C., Dart, E. H., & Wright, S. J. (2018). The effect of data points per x- to y-axis ratio on visual analysts' evaluation of single-case graphs. *School Psychology Quarterly, 33*(2), 314–322. https://psycnet.apa.org/doi/10.1037/spq0000243

Sommet, N., & Morselli, D. (2017). Keep calm and learn multilevel logistic modeling: A simplified three-step procedure using Stata, R. Mplus, and SPSS. *International Review of Social Psychology, 30*(1), 203–218. https://doi.org/10.5334/irsp.90

van den Bosch, R. M., Espin, C. A., Chung, S., & Saab, N. (2017). Data-based decision-making: Teachers' comprehension of curriculum-based measurement progress-monitoring graphs. *Learning Disabilities Research & Practice*, *32*(1), 46–60. https://doi.org/10.1111/ldrp.12122

Wagner, D. L., Hammerschmidt-Snidarich, S. M., 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. https://doi.org/10.1111/ldrp.12125