Journal of School Psychology 58 (2016) 41–55

Contents lists available at ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc

How accurate are interpretations of curriculum-based measurement progress monitoring data? Visual analysis versus decision rules☆

Ethan R. Van Norman [a,*], Theodore J. Christ [b]

[a] Georgia State University, United States
[b] University of Minnesota, Twin Cities, United States

ARTICLE INFO

Article history:
Received 23 July 2015
Received in revised form 25 April 2016
Accepted 15 July 2016
Available online xxxx

Keywords:
Assessment
Curriculum based measurement
Progress monitoring

ABSTRACT

Curriculum based measurement of oral reading (CBM-R) is used to monitor the effects of academic interventions for individual students. Decisions to continue, modify, or terminate these interventions are made by interpreting time series CBM-R data. Such interpretation is founded upon visual analysis or the application of decision rules. The purpose of this study was to compare the accuracy of visual analysis and decision rules. Visual analysts interpreted 108 CBM-R progress monitoring graphs one of three ways: (a) without graphic aids, (b) with a goal line, or (c) with a goal line and a trend line. Graphs differed along three dimensions, including trend magnitude, variability of observations, and duration of data collection. Automated trend line and data point decision rules were also applied to each graph. Inferential analyses permitted the estimation of the probability of a correct decision (i.e., the student is improving – continue the intervention, or the student is not improving – discontinue the intervention) for each evaluation method as a function of trend magnitude, variability of observations, and duration of data collection. All evaluation methods performed better when students made adequate progress. Visual analysis and decision rules performed similarly when observations were less variable. Results suggest that educators should collect data for more than six weeks, take steps to control measurement error, and visually analyze graphs when data are variable. Implications for practice and research are discussed.

1. Introduction

Within multi-tiered systems of support (MTSS), students who receive supplemental instruction are periodically assessed to determine whether they are benefiting appropriately from a given intervention (Fuchs & Stecker, 2003). Researchers often cite curriculum-based measurement of oral reading (CBM-R) as a tool especially suited to formatively assess reading instruction across relatively brief periods of time (Deno, Marston, & Tindal, 1986). When educators use CBM-R, they administer a grade level passage of connected text and calculate the number of words read correctly in one minute (WRCM; Deno, 1985). Educators then administer alternate forms across time, plot the WRCM scores on time series graphs, and evaluate the resulting pattern of observations to make decisions to continue or modify interventions (Deno, 1986). Generally, these decisions are made by visually analyzing data or applying decision rules in the context of a single-case design.

http://dx.doi.org/10.1016/j.jsp.2016.07.003

1.1. Visual analysis

Single-case designs enable researchers and educators to identify evidence based interventions in applied settings with individual or small groups of students. Effective interventions are primarily identified by establishing a functional relationship between a target behavior and an intervention. To establish a functional relationship, behavior is repeatedly measured in the absence (baseline phase) and presence of an intervention (treatment phase; Sidman, 1960). As observations are collected, clinicians conduct visual analysis to interpret differences in level (average performance), trend (upward or downward pattern of observations), and variability (spread of observations) across phases (Kazdin, 2011). Depending on the behavior, a change in level, trend, or variability across phases may be indicative of a functional relationship between an intervention and target behavior. Though a variety of strategies have been proposed, visual analysis remains the predominant method to interpret treatment effects within single-case designs (Ximenes, Manolov, Solana, & Quera, 2009).

1.2. Visual analysis and CBM-R progress monitoring

Certain characteristics of CBM-R, primarily the ability to repeatedly measure student behavior across time, makes the tool particularly amenable for use within single-case designs (Deno, 1986, 2003). With that said, the process of visually analyzing CBM-R progress monitoring data in schools is distinct from that associated with traditional single-case designs. When collecting CBM-R data in schools, educators rarely alternate between baseline and intervention phases; instead practitioners often use a continuous intervention design (Van Norman, Nelson, Shin, & Christ, 2013). A student is assumed to be making adequate progress when a steep upward pattern of observations with minimal variability is observed. Though such designs do not permit conclusive statements regarding functional relationships (thus making them unacceptable in academic intervention research; Gast, 2010; Kazdin, 2011), they do support conclusions regarding whether a student is improving in the presence of a given instructional program. The present discussion of visual analysis and CBM-R reflects decision making practices using a continuous intervention design with individual students.

Researchers have begun to evaluate the accuracy of visual analysts' judgements regarding CBM-R data. For instance, Van Norman et al. (2013) found that trend magnitude influenced the probability of correctly identifying whether a student was responding to an intervention among visual analysts who interpreted CBM-R progress monitoring data. The authors of that study simulated a continuous intervention design where one CBM-R observation was collected per week for 10 weeks. The probability of a correct decision (i.e., accurate identification of whether or not the student was responding to instruction) was much higher when a progress monitoring case depicted a substantial weekly rate of improvement (ROI; 3.00 WRCM per week) or no improvement (0.00 WRCM per week). The probability of a correct decision was lowest when ROI was minimal (0.75 WRCM per week). In addition, the authors found that graphic aids, such as goal lines and trend lines, increased the probability of a correct decision across all trend magnitudes. Within the study the researchers did not manipulate the duration of data collection or the variability of observations.

1.3. Decision rules

In the interest of not relying upon potentially subjective visual judgments of CBM-R data, some educators employ formal rules to guide decisions to modify instructional programs. In the context of CBM-R, such rules compare a student's ROI to an expected ROI based upon a meaningful long-term goal (i.e., aim line or goal line). The slope of the goal line quantifies the weekly gain in WRCM the student needs to reach a long-term goal. In their systematic review of CBM-R decision rules, Ardoin, Christ, Morena, Cormier, and Klingbeil (2013) identified over 100 documents that offered suggestions on how to apply data point and trend line decision rules (see below for more information regarding each rule type). The most pertinent finding for the present study was that through 2010, no empirical investigations evaluated the accuracy of either rule.

1.4. Data point rules

Data point decision rules originated from the precision teaching literature (White & Haring, 1980). When using a data point rule, treatment decisions are based upon the most recent observations. For instance, if the last three observations fall below the goal line, an instructional modification is considered. If the three most recent observations fall above the goal line, the slope of the goal line is increased (Fuchs, Fuchs, & Hamlett, 1989). If the three most recent observations are distributed both above and below the goal line, the current instructional program is maintained (Cates & Kitkowski, 2010). There are variations of the rule, which rely upon evaluating the four of five most recent consecutive data points to guide decisions (Fuchs & Fuchs, 2007).

To evaluate the accuracy of recommendations from data point decision rules, Van Norman and Christ (in press) used probability theory and a spreadsheet program to derive the likelihood of making a correct decision when using a three-point rule relative to a fixed 1.50 WRCM goal line. The authors assumed that one observation was collected per week and manipulated the magnitude of true growth, the number of weeks of progress of monitoring, and the variability in data (i.e., spread of observations around a line of best fit). They found that the probability of correctly identifying whether a student was responding to instruction never exceeded chance when true growth approximated the goal line. Conversely, when true growth was considerably greater than or less than the goal line, correct decisions could be made after about 11–12 weeks with somewhat variable data.