Studies in Educational Evaluation 59 (2018) 67–75

Contents lists available at ScienceDirect

Studies in Educational Evaluation

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

Biased predictions of students' future achievement: An experimental study on pre-service teachers' interpretation of curriculum-based measurement graphs

Florian Klapproth

Medical School Berlin, Calandrellistrasse 1-9, 12247 Berlin, Germany

ARTICLE INFO

Keywords:
Curriculum-based measurement
Experiment
Reading progress
Prediction

ABSTRACT

In an online experiment, a sample of $N = 109$ pre-service teachers were presented with 14 graphs mimicking graphs used in curriculum-based measurement. Graphs depicted a student's weekly test scores for the first part of a semester, and participants were instructed to use the graphs to predict students' achievement at the end of the semester. Relative to a linear regression model, participants generally tended to underestimate future achievement (i.e., predictions were negatively biased). Predictions were more negatively biased when data variability was low rather than high, when improvement was steep rather than flat, and when the most recent score indicated a performance upturn as opposed to downturn. The results are interpreted in the light of models of judgmental anchoring (Kahneman & Tversky, 1973; Mussweiler & Strack, 1999). Implications for practice are discussed.

1. Introduction

Curriculum-based measurement (CBM) is an increasingly popular method teachers can use to track how students are progressing in basic academic areas such as math, reading, spelling or writing. CBM entails using quick, frequently-administered, standardized measures to assess students' progress towards a long-term goal (Deno, 1985). The primary purpose of such repeated assessment is to provide teachers with a meaningful basis for evaluating the success of instruction (Deno, 1986) and for identifying students in need of additional support or modified instruction (Fuchs, Fuchs, McMaster, & Al Otaiba, 2003; Hosp, Hosp, & Howell, 2007). Although CBM offers educators a sound foundation for making evidence-based decisions, the effectiveness of CBM for improving student achievement appears to be mixed, and frequent progress monitoring alone does not appear to improve student achievement (Stecker, Fuchs, & Fuchs, 2005).

One reason why CBM alone does not lead to better student achievement is that teachers may have difficulties using CBM data to make reasonable predictions about students' future achievement (Van den Bosch, Espin, Chung, & Saab, 2017). Although it is well known that people are subject to a number of cognitive biases when making predictions (Lawrence, Goodwin, O'Connor, & Önkal, 2006), to date little is known about systematic bias in how educators use CBM data to predict students' future achievement. Given that predictions of student achievement are important for (1) deciding whether or not to change instruction (Good & Shinn, 1990), (2) establishing long-term goals in education (Good & Shinn, 1990; Shinn, Good, & Stein, 1989), and (3) identifying students at risk (Hosp et al., 2007; Strathmann, Klauer, & Greisbach, 2010), it is critical to investigate the conditions under which CBM data might lead educators to make biased predictions. The current study therefore uses an experiment to examine how the variability of scores, the rate of improvement, and a student's last, most recent score, each affect educators' expectations for how that student will perform in the future.

1.1. Using CBM data to predict future achievement

With CBM, teachers evaluate a student's progress based on the set of observations made so far (e.g., the first half of a school semester or year), and then *predict* whether, given the observed rate of improvement, a student is "on track" to meet a particular educational goal within a certain time frame (e.g., by the end of the semester or year) (Good & Shinn, 1990; Van Der Heyden, Snyder, Broussard, & Ramsdell, 2008). The slope of the available progress data summarizes a student's current rate of improvement, and also represents a reasonable prediction about how well that student will perform in the future if instruction remains unchanged (Shinn, Good et al., 1989). Thus, with CBM, educators must first estimate the rate of improvement inherent in the available data, and then extrapolate whether, given the observed trend, the student is likely to meet a performance standard by a specific time

E-mail address: florian.klapproth@medicalschool-berlin.de.

https://doi.org/10.1016/j.stueduc.2018.03.004
Received 1 September 2017; Received in revised form 14 March 2018; Accepted 21 March 2018

point. Indeed, such "trend line" decision rules are commonly used to help educators decide whether an intervention has resulted in adequate student progress, and therefore whether the intervention should be continued, modified, or abandoned (Ardoin, Christ, Morena, Cormier, & Klingbeil, 2013).

Usually, researchers and practitioners recommend regression techniques to estimate the trend and to predict students' future achievement (Shinn, Good et al., 1989; Strathmann et al., 2010). In practice, however, educators do not always interpret CBM graphs appropriately (Stecker et al., 2005), and there are several reasons why educators may fail to use sound statistical methods to interpret the data (Van Norman, Nelson, Shin, & Christ, 2013). Educators may lack access to analytical software (Shinn, Good et al., 1989), or they may feel uncomfortable using or mistrust computer programs designed to analyze and interpret CBM data (e.g., Fuchs & Fuchs, 1989; Landrum, Cook, Tankersley, & Fitzgerald, 2007). Moreover, computer applications can distance educators from CBM data and thereby limit the extent to which they are willing and able to use and meaningfully interpret the data (Fuchs, Fuchs, & Stecker, 1989). Thus, educators often visually judge student progress and predict future achievement as opposed to following data-driven decision rules and statistical procedures (Van Norman et al., 2013). Even when they have support from computer software, educators should nevertheless be able to independently interpret CBM data and make sound data-based predictions (Van den Bosch et al., 2017). For instance, CBM data patterns can be ambiguous, and educators need to be able to correctly recognize the ambiguity inherent in the data in order to make the best decision for a particular student (Deno, 2013). Ambiguity can occur when the features of the graph do not consistently speak in favour or against the existence of a positive or negative trend, for instance, when data cyclicity is present (Brossart, Parker, Olson, & Mahadevan, 2006).

In Germany, the use of CBM both in regular school and in special education has become more frequent, as is indicated by different tools available for teachers to formatively evaluate their instruction (e.g., Diehl & Hartke, 2012; Souvignier, Förster, & Salaschek, 2014; Strathmann & Klauer, 2010; Walter, 2013; Wilbert, 2014). For instance, a comprehensive and computer-assisted tool ("quop") for examining competences in mathematics, reading and second language has been developed by Souvignier et al. (2014). This tool provides teachers with graphs visually representing the development of the students' skills, and furthermore with statistics, e.g. the slope of the trend. However, teachers seem to prefer visual inspection of the data instead of using statistics when analysing and reporting the data (Jain & Spieß, 2012).

1.2. Cognitive processes related to CBM

According to Tversky's and Kahneman's anchor-adjust heuristic model (Hogarth & Makridakis, 1981; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974), people use "anchors" to make an estimate or a prediction. As an example of anchoring, one study indicated that people's estimates of the percentage of African countries represented in the United Nations largely depended on an arbitrarily selected number (Tversky & Kahneman, 1974). People are often aware that an anchor is inaccurate and may adjust their initial estimate (i.e., the anchor) in different ways and to different degrees. For instance, there is some evidence that people use the last point of a time series as an anchor, and adjust the anchor towards the mean of the series (cf. Bolger & Harvey, 1993; Eggleton, 1982; Harvey, 2007). This would explain why people tend to underestimate future values when the available data suggest a positive trend (Bolger & Harvey, 1993; Eggleton, 1982; Lawrence & Makridakis, 1989; Sanders, 1992), whereas people tend to overestimate future values when the trend is negative (Wagenaar & Sagaria, 1975).

An elaboration of the anchor-adjust heuristic is the selective accessibility model proposed by Mussweiler and Strack (e.g., Mussweiler & Strack, 1999). According to this model, people relate the to be estimated value to a standard value and selectively search for information that is consistent with the standard. This selective accessibility mechanism leads to an assimilation effect, meaning that predictions will be similar to the standard. For instance, Mussweiler and Schneller (2003) presented their participants with graphs indicating past stock prices, with either extreme high or extreme low past prices within the price development. They found that expectations about future prices assimilated to the extreme past prices.

Since in CBM teachers rely their decisions frequently on time-series data by visual inspection, both the anchor-adjust heuristic and the selective accessibility model would anticipate that predictions of students' future achievement would assimilate towards either the mean (anchor-adjust heuristic) or the extreme values of the series (selective accessibility model). Both assimilation effects would result in negative bias, meaning an underestimation of the trend, if the trend is positive (anchor-adjust) or the extreme values are rather low (selective accessibility model).

Predictions based on the graphical presentation of time-series data may also be affected by the number of spatial transformations that are necessary to mentally extend the trend line to future time points (Trickett & Trafton, 2004, 2006) The more spatial transformations have to be conducted, i.e., the more the trend line has to be extended, the less accurate would be the prediction (Trickett & Trafton, 2004).

1.3. Characteristics of CBM data may lead to bias

Given the complex cognitive processes involved in interpreting data and making predictions, it is unsurprising that educators have difficulties interpreting CBM graphs (Friel, Curcio, & Bright, 2001; Glazer, 2011; Van den Bosch et al., 2017). Furthermore, it is well known that human judges are prone to making a number of cognitive errors, and that their predictions are likely to be biased in the absence of more objective procedures (e.g., statistical analyses; cf. Lawrence et al., 2006).

Certain characteristics of student progress data might increase the chance of biased predictions. For instance, achievement scores are often contaminated by construct-irrelevant variance (Christ, Zopluoglu, Long, & Monaghen, 2012), which may make it difficult to visually identify systematic patterns in the data. With CBM progress data in particular, people have difficulty accurately estimating the rate of improvement when progress data are highly variable (Nelson, Van Norman, & Christ, 2017; Tindal, Deno, & Ysseldyke, 1983) or include extreme values (Nelson et al., 2017). In the Nelson et al. (2017) study, participants had more difficulties detecting a negative trend as opposed to a positive trend when data variability was high. Hence, they tended to overestimate student progress (and hence, students' future achievement) when variability was high.

People's ability to accurately estimate a data trend also seems to depend on the direction of the trend. Studies on time-series data in general have shown that participants tend to underestimate positive linear trends (Eggleton, 1982; Lawrence & Makridakis, 1989; Klapproth, 2006; Wagenaar & Sagaria, 1975; Wagenaar & Timmers, 1978). Furthermore, the degree of underestimation seems to depend on the rate of improvement, with steeper slopes corresponding with stronger underestimation (e.g., Sanders, 1992). That is, people usually do not "believe" that a strong positive trend (e.g., a student's high rate of improvement) will continue at the same rate.

In addition to variability and the rate of improvement, some studies have shown that the most recent observation of a time series disproportionately affects predictions (Harvey, Bolger, & McClelland, 1994; Reimers & Harvey, 2011). Because people tend to anchor future predictions around the most recent observation, a student's most recent test score may likewise disproportionately affect educators' predictions. Presumably, the most recent test score should affect predictions to a greater extent when test scores are highly variable, because changes in a series are more likely to be judged as indicating a trend in more "noisy" data series (Goodwin & Wright, 1993).