The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2397-7604.htm

Data-driven decision-making in creating class rosters

Rebecca Wolf, Joseph M. Reilly and Steven M. Ross
Center for Research and Reform in Education, Johns Hopkins University, Baltimore, Maryland, USA

Received 19 March 2019
Revised 21 June 2019
16 October 2019
Accepted 23 December 2019

Abstract

Purpose – This article informs school leaders and staffs about existing research findings on the use of data-driven decision-making in creating class rosters. Given that teachers are the most important school-based educational resource, decisions regarding the assignment of students to particular classes and teachers are highly impactful for student learning. Classroom compositions of peers can also influence student learning.

Design/methodology/approach – A literature review was conducted on the use of data-driven decision-making in the rostering process. The review addressed the merits of using various quantitative metrics in the rostering process.

Findings – Findings revealed that, despite often being purposeful about rostering, school leaders and staffs have generally not engaged in data-driven decision-making in creating class rosters. Using data-driven rostering may have benefits, such as limiting the questionable practice of assigning the least effective teachers in the school to the youngest or lowest performing students. School leaders and staffs may also work to minimize negative peer effects due to concentrating low-achieving, low-income, or disruptive students in any one class. Any data-driven system used in rostering, however, would need to be adequately complex to account for multiple influences on student learning. Based on the research reviewed, quantitative data alone may not be sufficient for effective rostering decisions.

Practical implications – Given the rich data available to school leaders and staffs, data-driven decision-making could inform rostering and contribute to more efficacious and equitable classroom assignments.

Originality/value – This article is the first to summarize relevant research across multiple bodies of literature on the opportunities for and challenges of using data-driven decision-making in creating class rosters.

Keywords Class rosters, Data-driven decisions, Teacher quality, Peer effects

Paper type Literature review

Introduction

In preparation for each school year, schools must perform a routine task that can have a substantial impact on student learning: creating class rosters, or determining which students are grouped into classes with which teachers. *Rostering* potentially has a substantial impact on student outcomes because teachers are the most important school-based factor impacting student learning (Rice, 2003; Sanders and Rivers, 1996). Despite these potential effects, however, school leaders and staffs often rely on subjective judgments in assigning students to classes and do not fully leverage the data they have at their disposal (Bosworth, 2014; Dieterle *et al.*, 2012; Kraemer *et al.*, 2011). Moreover, these subjective judgments can lead to inequitable teacher-student assignments (Dieterle *et al.*, 2012; Kalogrides and Loeb, 2013).

The lack of data-driven decision-making in rostering is not surprising. Rostering is complex, and school staff may consider a host of variables including teacher effectiveness, student achievement, student demographic characteristics and others (Burns and Mason, 2002; Henderson, 2011; Kraemer *et al.*, 2011). Therefore, creating class assignments can be

© Rebecca Wolf, Joseph M. Reilly and Steven M. Ross. Published in *Journal of Research in Innovative Teaching & Learning*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/4.0/legalcode

Journal of Research in Innovative Teaching & Learning
Vol. 14 No. 2, 2021
pp. 162-177
Emerald Publishing Limited
2397-7604
DOI 10.1108/JRIT-03-2019-0045

both logistically demanding and time-consuming for school staff (Burns and Mason, 2002; Hopkins, 1999). According to Henderson (2011), examining how data can be systematically incorporated into the rostering process is an unrealized practice that could be beneficial for school leaders and staffs.

Do data exist that could improve rostering in ways that would enhance student achievement and other educational outcomes? Data have become both increasingly available over time, as educational policy at the national, state and district levels has emphasized the use of data-driven decision-making (Anderson *et al.*, 2010; Cannata *et al.*, 2017; Coburn and Turner, 2011; Marsh, 2012). Federal initiatives have also incentivized the collection and use of data on teacher effectiveness (Coburn and Turner, 2012; Donaldson and Papay, 2015; Drake *et al.*, 2016; Harris and Herrington, 2015; Master, 2014). In addition to data on teacher effectiveness, schools also have access to student achievement, disciplinary and demographic data. The question is, can school leaders and staffs use these data to optimize classroom assignments, and if so, how? This review addresses this overarching question and provides school leaders and staffs responsible for creating class rosters with information that can help them improve the equity and efficacy of their classroom assignments.

What is data-driven decision-making?

It is important to discuss how data-driven decision-making was conceptualized for the purposes of this study. Data-driven decision-making begins with the existence of data. Data generally refer to existing information that can be systematically analyzed in such a way to provide new information to practitioners (Coburn and Turner, 2011; Ikemoto and Marsh, 2007; Marsh, 2012). Therefore, in this review, data were limited to quantitative metrics that are routinely collected by districts and schools or can be derived from routinely collected data.

Another key consideration in data-driven decision-making is that quantitative data are not necessarily interpreted the same way by different stakeholders. Practitioners make sense of new information using their prior ideologies, beliefs, experiences and expertise (Coburn and Turner, 2011; Ikemoto and Marsh, 2007; Marsh, 2012; Park *et al.*, 2012). Therefore, different people may come to different conclusions when presented with the same information (Coburn, 2010; Coburn and Turner, 2011; Ikemoto and Marsh, 2007; Marsh, 2012). Some practitioners, in fact, may discount new information that they do not understand, that challenges their beliefs, or that "raises questions about the efficacy of past practices" (Coburn and Turner, 2011, p. 179). As Spillane (2012) stated, "Data do not objectively guide decisions on their own—people do" (p. 114).

Once new information has been processed into new knowledge, practitioners must decide whether to act on the new knowledge, and their actions reflect data-driven decision-making (Ikemoto and Marsh, 2007; Marsh, 2012). After taking action, practitioners may enter a cycle of continuous improvement, where new data are collected to monitor progress, and the new data will once again be processed into new information, which could inform future actions (Marsh, 2012). Successful data-driven decision-making results in improved educational offerings and outcomes (Coburn and Turner, 2011, 2012; Ikemoto and Marsh, 2007; Park *et al.*, 2012; Slavin *et al.*, 2013).

The success of data-driven decision-making depends on multiple factors. Data must be accessible, timely, and valid or perceived as such (Coburn and Turner, 2011; Ikemoto and Marsh, 2007; Marsh, 2012). Practitioners must have the willingness and capacity to analyze the data and process the information into new knowledge (Coburn and Turner, 2011; Ikemoto and Marsh, 2007; Marsh, 2012). Practitioners must also have the autonomy to make decisions based on the new knowledge generated from the data.

Data-driven decision-making works best in trusting environments, where practitioners are not deterred from asking hard questions (Coburn and Turner, 2011; Ikemoto and Marsh, 2007; Marsh, 2012). Trust is also needed because data-driven decision-making rarely occurs with a single individual and is more likely to be collective in nature (Coburn and Turner, 2011), particularly when it comes to rostering (Burns and Mason, 2002; Cohen-Vogel, 2011; Hopkins, 1999). Therefore, it is critical that the practitioners engaged in collective data-driven decision-making perceive the process as constructive rather than punitive (Ingram *et al.*, 2004). Moreover, when data-driven decision-making is limited to a single individual, turnover in school staff can substantially hinder data use (Coburn and Turner, 2011; Marsh *et al.*, 2008).

School leaders are ideal candidates to create trusting environments in which to process data and ultimately establish a school-wide culture for data use (Coburn and Turner, 2011; Ikemoto and Marsh, 2007; Park *et al.*, 2012). School leaders can frame how new information is presented to teachers and staff to garner more buy-in and potentially challenge “long-standing norms, beliefs, and structure (e.g. about student ability and tracking) that may inhibit inquiry” (Coburn and Turner, 2011, p. 667). Norms, beliefs and structures have implications for data-driven decision-making in creating class rosters, as teachers may prefer to teach some subgroups of students but not others (Cannata, 2010; Carey *et al.*, 1994; Engel *et al.*, 2014; Kalogrides *et al.*, 2013).

Yet, little is known about how school leaders and staffs use data to inform the rostering process (Henderson, 2011). This literature review sheds light on how school leaders and staffs have used data-driven decision-making in the rostering process in the past and how they might use existing data to improve the process moving forward. To the authors’ knowledge, this article is the first review of research on topics relevant to rostering with the goal of informing how data-driven decision-making may be used to improve the rostering process. Specifically, this review addresses the following research questions:

RQ1. How have school leaders and staffs used data in the rostering process?

RQ2. What were systematic patterns in teacher–student assignments?

RQ3. How were measures of teacher effectiveness used in the rostering process?

RQ4. What are considerations for using other teacher characteristics in the rostering process?

RQ5. Should peer effects be considered in the rostering process?

Method

To conduct the literature review on the use of data-driven decision-making in the rostering process, databases were searched for studies using keywords and their synonyms related to class rosters and assignments, teacher–student assignments, peer effects and teacher effectiveness. Databases searched included Google Scholar, EBSCOhost, ProQuest Central, and the US Department of Education’s Education Resources Information Center database. Following these initial database searches, the reference lists of particularly relevant articles were mined to identify additional studies that were either not initially identified through the database search or that predated the search criteria. This procedure included identifying the key authors who were cited frequently in many of the studies and conducting additional searches targeting the works of these researchers. Finally, secondary sources and meta-analyses on topics related to rostering were reviewed in order to identify how research on this topic had been synthesized thus far. Studies were also limited to those conducted in the United States, and with rare exception, to those published between 2007 and 2017. This process ultimately generated a total of 265 articles that were initially identified and reviewed by researchers.

Of the 265 studies initially reviewed, 58 studies directly addressed the research questions, and findings from these studies are synthesized in the subsequent section. Per this study’s conceptualization of data-driven decision-making, the review targeted studies that focused on the use of quantitative data — as opposed to qualitative data — in rostering decisions.

After studies were identified, findings were summarized using a narrative overview approach (Green *et al.*, 2006). This approach allows for researchers to summarize across research findings and different bodies of literature (Green *et al.*, 2006). The narrative overview approach was viewed as appropriate given the exploratory nature of this study and the many bodies of literature that either directly or indirectly related to the rostering process.

Findings

The following sections summarize the findings from the literature review about how school leaders and staffs have used data to inform the rostering process and how existing administrative data may be used to improve classroom assignments.

How have school leaders and staffs used data in the rostering process?

A few studies documented how school leaders have used data in the rostering process (Burns and Mason, 2002; Henderson, 2011; Kraemer *et al.*, 2011), yet prior research is limited. According to these studies, school staffs considered a number of data elements in the rostering process. Henderson (2011) surveyed principals across roughly 30 elementary schools in a single North Carolina school district, and Kraemer *et al.* (2011) conducted interviews and focus groups with principals from roughly 30 K–8 schools across three large urban districts. Findings across the two studies showed that school staffs frequently used the following data elements in the rostering process (Henderson, 2011; Kraemer *et al.*, 2011):

(1) **Student academic performance**: Student test scores.

(2) **Student demographic characteristics**: Student gender, race/ethnicity, primary language, socioeconomic status, English learner (EL) status, and special education status.

(3) **Other student characteristics**: Gifted and talented status, participation in elective coursework, grade-level retention and disciplinary incidents.

(4) **Teacher characteristics**: Measures of teacher effectiveness, years of experience and certification type or status.

Other non-quantitative metrics considered in the rostering process (Henderson, 2011; Kraemer *et al.*, 2011) were as follows:

(1) **Student attributes**: Subjective judgments made by a student's previous teacher about a student's personality, learning style, parent characteristics and disruptive behavior or misbehavior related to specific peers.

(2) **Teacher attributes**: Administrator perceptions of a teacher's ability to effectively teach students at different grade and ability levels, and of the teacher's instructional and classroom management styles and capabilities.

(3) **School structures**: Team teaching and teacher "looping" (remaining with the same students as they progress to the next year).

Findings indicated that to some degree, rostering decisions were made intentionally (Burns and Mason, 2002; Henderson, 2011; Kraemer *et al.*, 2011). School staffs frequently cited creating "balanced" classrooms across multiple variables as a major consideration (Burns and Mason, 2002; Kraemer *et al.*, 2011, p. 5). A frequent interest, for example, was increasing similarities between classrooms in student achievement and characteristics (Burns and Mason, 2002; Henderson, 2011; Kraemer *et al.*, 2011). One study found that school staffs often attempted to achieve balance by first assigning special education and EL students, who required specialized academic programming, to classrooms, and then building rosters around them (Burns and Mason, 2002). Another study reported that, "within the balancing structure," school staffs matched specific students and teachers on the basis of their characteristics or "interaction tendencies" with other students (Kraemer *et al.*, 2011, p. 1).

Henderson (2011) also analyzed the degree to which various rostering practices were correlated with student growth in reading. This research found that considerations of teacher grade-specific years of experience and the strategic placement of high-achieving students were each positively correlated with the schools' average growth in reading. Other studies have corroborated the finding that grade switching of teachers between years had at least a short-term negative effect on teacher effectiveness, particularly for novice teachers (Atteberry *et al.*, 2016; Blazar, 2015; Ost, 2014). On the other hand, Henderson (2011) determined that subjective considerations of parents and student learning styles were each negatively correlated with schools' average reading growth. Other variables considered in the rostering process (and included in the list above) were either unrelated or weakly correlated with school-level average reading growth. While Henderson's (2011) findings are correlational and not causal, they suggest that leveraging data-driven decision-making in the rostering process could be impactful for student achievement.

Prior studies have also found that in using data to inform rostering, school staffs tended to rely on paper-based forms of data, as opposed to electronic systems. One study gave the example of school staffs creating information cards for each student to be used in the rostering process (Burns and Mason, 2002; Kraemer *et al.*, 2011). Kraemer and colleagues (2011) also found that data used for rostering were typically in paper form, and school staffs seldom used electronic systems to create rosters. High schools and some large middle schools that adopted high school scheduling models were the exceptions. The reliance on paper-based forms of data may impede data-driven decision-making in creating class rosters, however. One study, for instance, found that school leaders and teachers were more likely to see the value in data-based decision-making when they had online access to multiple data points and could disaggregate the data and display the results in more than one way, although not all districts offered this capability (Kerr *et al.*, 2006). In sum, prior research indicates that school staffs have used data, at least in some format, in the rostering process, although there is less evidence that school staffs have engaged in data-based decision-making in systematic ways.

What were systematic patterns in teacher–student assignments?

The lack of data-driven decision-making in creating class rosters has potential drawbacks. While some studies reported that many school staffs attempted to create "balanced" classes, other studies indicated that whether intentionally or unintentionally, classroom rosters were often not balanced, particularly in terms of student prior achievement (Dieterle *et al.*, 2012; Kalogrides and Loeb, 2013). Even in the absence of formal tracking, students were often grouped in classrooms by prior achievement, a finding evidenced in elementary, middle and high schools (Dieterle *et al.*, 2012; Kalogrides and Loeb, 2013).

Researchers have long argued over whether grouping students within classrooms by prior ability impacts student learning. Although individual studies of ability grouping have produced mixed results, meta-analyses of many studies have found no overall effect (Slavin, 1990; Steenbergen-Hu *et al.*, 2016). Mixed results across individual studies are somewhat expected, however, because it is difficult to disentangle the effects of ability grouping from other phenomena.

Access to effective teachers, for example, may be mediated by classroom assignments (Clotfelter *et al.*, 2006; Ingersoll, 1999; Kelly, 2004; Player, 2010). In fact, prior studies have consistently found that more effective teachers were disproportionately assigned to higher performing students, while less effective teachers were disproportionately assigned to lower performing students (Aaronson *et al.*, 2007; Kalogrides *et al.*, 2013; Rothstein, 2009). This body of literature defined teacher effectiveness in terms of value-added measurements (VAMs), or the degree to which improvements in students' test scores over time could be attributed to the teacher using statistical modeling. Moreover, given the complex relationship between student performance, socioeconomic status, and race/ethnicity, studies have also found that low-income and historically disadvantaged minority students were disproportionately assigned to less effective teachers within a school than their peers (Bacher-Hicks *et al.*, 2014; Feng, 2010; Player, 2010).

The systematic assignment of more effective teachers to higher performing students and less effective teachers to lower performing students within a school may be the result of several factors. First, school leaders may attempt to reward or retain effective teachers by assigning them higher numbers of high-achieving or better behaved students (Player, 2010). Second, experienced teachers have generally been found to be more effective than novice teachers, and experienced teachers may have greater autonomy or status to either create or influence class rosters; not surprisingly, experienced teachers have often preferred to teach higher achieving students (Burns and Mason, 2002; Clotfelter *et al.*, 2006; Hanushek *et al.*, 2005; Hopkins, 1999; Jackson, 2014; Kraemer *et al.*, 2011).

As teachers are the most important school-based resource impacting student learning (Rice, 2003), the practice of assigning more effective teachers to higher performing students is problematic to the extent that it yields inequitable access to effective teachers. This practice may also result in other undesirable outcomes. For example, Feng (2010) found that novice teachers were more likely to have tougher classroom assignments, and tougher classroom assignments were associated with increased likelihoods that novice teachers left the school or profession. Tougher classroom assignments, however, had weaker effects for more experienced teachers. Using data-driven decision-making therefore provides a potential means to improve the equity of the rostering process for both students and teachers, relative to more subjective and informal approaches.

How were measures of teacher effectiveness used in the rostering process?

There is some evidence that elementary school leaders have used measures of teacher effectiveness in the rostering process, although not for the purpose of ensuring equity in teacher–student assignments. Several studies found school leaders used measures of teacher effectiveness in terms of VAMs to reassign ineffective teachers out of tested grades and into non-tested grades (Chingos and West, 2011; Cohen-Vogel, 2011; Fuller and Ladd, 2013; Goldring *et al.*, 2015). Notably, this practice has proven more prevalent than the converse strategy of moving the most effective teachers from untested into tested grades (Cohen-Vogel, 2011; Fuller and Ladd, 2013). Although the goal of reassigning ineffective teachers out of tested grades was undoubtedly to increase student performance on assessments used for school accountability, reassigning the least effective teachers to the non-tested and, by default, the youngest grades in elementary school, has obvious drawbacks as students build upon their academic knowledge as they progress from one grade to the next.

School leaders could theoretically consider teacher effectiveness in the rostering process to ensure equitable access to effective teachers, yet there are several potential challenges to doing so. If teacher effectiveness is measured in terms of VAMs, school leaders may not understand or trust VAMs and therefore may not perceive them as valid measures (Goldring *et al.*, 2015). Another issue is the timeliness of VAM data, which typically do not become available to school leaders until after the start of the school year when class rosters have already been established. Although school leaders could use prior years' VAM data, most appear to be reluctant to do so (Goldring *et al.*, 2015). Perhaps most problematic is that VAM data may be missing for a large proportion of teachers (Jiang *et al.*, 2015; Master, 2014).

Teacher effectiveness may also be measured on the basis of classroom observations. Principal ratings derived from classroom observations are the most common teacher evaluation method and also assess important teacher skills that are not captured by VAMs (Goe *et al.*, 2008; Harris and Sass, 2014). A potential drawback to relying on principal ratings, however, is that these ratings may be more subject to bias, and consequently, less predictive of student learning gains than those of external, non-peer reviewers (Harris and Sass, 2014; Measures of Effective Teaching (MET) Project, 2013; Whitehurst *et al.*, 2014). Yet the degree of bias appears to decrease as the number of years the principal and teacher know each other increases (Rockoff *et al.*, 2012). In other words, principals appear to more accurately rate teachers after they have observed the teacher over multiple years.

Taken together, this body of research indicates that consideration of teacher effectiveness in the rostering process could serve some utility in ensuring equitable access to effective teachers. However, caution is needed to ensure that the measures of teacher effectiveness reflect true differences in teacher quality, as opposed to rater bias.

What are considerations for using other teacher characteristics in the rostering process?
An important consideration in the rostering process is whether teachers are differentially effective with some groups of students. For example, are some teachers more effective with low-achieving vs high-achieving students? A growing body of research has explored this question using VAMs (Fox, 2016; Harris, 2009; Loeb and Candelaria, 2012; Papay, 2011; Reardon and Raudenbush, 2009). In a leading study, Lockwood and McCaffrey (2009) determined that effective teachers were generally effective with all students, yet 10 percent of the total teacher effect could be attributed to being particularly effective with students at a certain ability level.

Researchers have also examined whether teachers were differentially effective with demographic subgroups of students. Study findings indicate that students may benefit if assigned to a same-race teacher (Aaronson *et al.*, 2007; Dee, 2004; Egalite *et al.*, 2015; Gershenson *et al.*, 2017; Yarnell and Bohrnstedt, 2018). Specifically, some studies found higher student learning gains for black students when assigned to a black teacher (Aaronson *et al.*, 2007; Dee, 2004; Egalite *et al.*, 2015; Yarnell and Bohrnstedt, 2018). One study even showed a reduced likelihood of high school dropout and persistence in a four-year college for low-income black students who had been assigned to at least one black teacher in the third through fifth grades (Gershenson *et al.*, 2017). The primary interpretation of these findings is that black teachers provide effective role models and are most sensitive to the backgrounds and needs of black students, particularly to male students (Yarnell and Bohrnstedt, 2018). However, such relationships are not necessarily straightforward, given Yarnell and Bohrnstedt's (2018) finding that black female (but not male) students equally benefitted from being assigned to Latino teachers. Studies have also estimated higher learning gains for white and Asian students when assigned to a same-race teacher, particularly for low-income students (Dee, 2004; Egalite *et al.*, 2015). Gender may interact with race as well. One study found that black females had higher learning gains when assigned to a female teacher (Aaronson *et al.*, 2007).

Prior studies have also explored whether teachers were differentially effective with ELs and native English speakers. Loeb and colleagues (2014) found that only 4 percent of teachers were simultaneously rated in both the top performance quintile for ELs and in the bottom quintile for non-ELs, or vice versa. The authors concluded that, in general, teachers who were effective with ELs were also effective with non-ELs. However, teachers who were fluent in Spanish and held bilingual certification were more effective for ELs than for non-ELs, on average. Therefore, while there may be modest benefit in strategically matching some students with some teachers, research suggests that effective teachers are generally effective with all student subgroups.

Finally, a few studies investigated whether elementary school teachers were differentially effective in teaching reading and mathematics (Condie *et al.*, 2014; Fox, 2016; Goldhaber *et al.*, 2013). For example, Fox (2016) in a simulation found that reassigning elementary school teachers to the subject of their strength could result in higher student learning gains. The study also found, however, that the teachers who were the most effective at teaching mathematics were also the most effective at teaching reading. Therefore, while assigning elementary school teachers to the subject of their strength may produce modest learning gains, caution is needed to avoid negative effects on gains in reading.

Should peer effects be considered in the rostering process?

Although teachers have a greater impact on student learning, research has also found that classroom peers can influence student learning (Burke and Sass, 2013; Figlio, 2007; Horoi and Ost, 2015; Hoxby and Weingarth, 2005; Imberman *et al.*, 2009; Lefgren, 2004; Liu, 2010; Sacerdote, 2011; Vigdor and Nechyba, 2007). Several studies have found negative effects on student learning when students were grouped with concentrations of low-achieving peers, and conversely, positive effects when students were grouped with concentrations of high-achieving peers (Burke and Sass, 2013; Domina *et al.*, 2016; Imberman *et al.*, 2009; Vigdor and Nechyba, 2007). For example, Imberman *et al.* (2009) analyzed the impact of Hurricane Katrina on student evacuees' achievement in grades 3–10 in Houston and found positive effects from exposure to high-achieving students and negative effects from exposure to low-achieving students. Similarly, Burke and Sass (2013) analyzed Florida data for students in grades 3–10 and found a positive relationship between improved student performance and average peer achievement in every grade and subject, with the exception of middle school mathematics. Moreover, these peer effects persisted over time (Burke and Sass, 2013). Similarly, using data from North Carolina, Hoxby and Weingarth (2005) found that exposure to high-achieving peers in grades 3–8 positively affected student learning. These studies also produced causal, as opposed to correlational, evidence of peer effects.

Peer effects appeared to be more complicated than a positive relationship between peer achievement and student performance, however. Burke and Sass (2013) found that low-achieving students benefitted from higher proportions of mid-achieving students in their classes. Mid-achieving students benefitted from higher proportions of high-achieving students in their classes but were hindered by higher proportions of low-achieving students. High-achieving students benefitted from higher proportions of either high- or low-achieving students in the class. The authors theorized that teachers may be able to more effectively target instruction when there is a narrower range of student abilities in a given class, and that students may learn better from peers whose ability level is within range of their own. To explain why high-achieving students may perform better with higher proportions of low-achieving students in their classroom, the authors hypothesized that high- and low-achieving students likely did not work together and thus high-achieving students reaped the benefit of collaborating with other high-achieving students. Hoxby and Weingarth (2005) similarly found that students benefitted from high-achieving peers when (1) there was at least a small proportion of like-ability peers in the class, and (2) there was not a wide disparity in peer ability in the class. These findings support the conclusion that peer achievement does impact student learning, but classroom dynamics are complex.

Peer behavior during class may also influence student learning. Using data from the National Longitudinal Study of Adolescent Health, Hill (2014) found that the probability of a student failing a middle or high school math course increased with the proportion of classmates who had previously failed and were repeating the course. Along those lines, studies have found that misbehaving peers can negatively impact student achievement (Figlio, 2007; Horoi and Ost, 2015; Sacerdote, 2011). Horoi and Ost (2015), for instance, analyzed the effect of being in a classroom with potentially disruptive peers (defined as students with emotional disabilities) and found that even the presence of one emotionally disabled student caused a decrease in academic performance for students in the class in both mathematics and reading. Similarly, Figlio (2007) found that the presence of disruptive classmates decreased mathematics achievement and increased the probability that the other classmates would exhibit behavior problems.

Correlational studies have found negative associations between concentrations of minority and low-income students and student learning (Berliner, 2009; Hanushek *et al.*, 2002; Kahlenberg, 2004; Liu, 2010; Rusk, 2002; Sacerdote, 2011; Van Ewijk and Sleegers, 2010). However, these studies cannot conclude that concentrations of these student subgroups *caused* lower achievement. Van Ewijk and Sleegers (2010) conducted a meta-analysis of socioeconomic peer effects at the classroom level and found negative associations between student learning gains and the proportion of low-income students in the class. They also found that students had higher average achievement when assigned to classes with higher proportions of socioeconomically affluent peers, regardless of the overall composition of students in their schools.

There are several potential reasons for negative peer effects associated with poverty. Compared with middle-class students, low-income students are more likely to have worse health, more school-to-school mobility, less-educated parents, a more limited vocabulary and fewer critical thinking skills (Rothstein, 2013). Moreover, to the extent that race is related to poverty, the same findings have occurred for some minority groups. Harris (2006) explained, "It is not race *per se* that affects learning, but the conditions under which minority students are raised and the characteristics of their classmates" (p. 18). It is important to note, however, that these empirical trends may be confounded with factors other than true poverty- or race-based peer effects. Researchers have not yet been able to successfully disentangle the extent to which these student characteristics were related to unmeasured factors, such as implicit bias of teachers (e.g. teachers who have large proportions of black students in their classes may unconsciously change their instructional practices and lower their standards) (Sacerdote, 2011). On the other hand, there is also limited evidence that Latino students performed better in classrooms when they were not racially isolated (Vigdor and Nechyba, 2007), and the same finding could extend for all minority racial groups.

Finally, studies have shown that higher proportions of female students in a class were associated with improved learning outcomes for both boys and girls (Hoxby, 2000; Lavy and Schlosser, 2007; Sacerdote, 2011). Lavy and Schlosser (2007) attributed this positive effect to greater proportions of female students in the class resulting in a better classroom and learning environment. Specifically, higher proportions of female students resulted in less misbehavior, better relationships among teachers and students, more teacher feedback and individualized instruction, and greater student and teacher satisfaction. In summary, these findings indicate that peer effects — in terms of achievement, behavior, socioeconomic status, race, and gender — may be consequential in student learning and therefore potentially worth consideration in the rostering process.

Discussion

Every year, schools across the country must assign students to particular teachers and classes. These seemingly routine decisions can have large effects on student learning, given that teachers are the most important school-based resource impacting student learning and that classroom peers can influence learning as well (Burke and Sass, 2013; Rice, 2003). This literature review synthesized research on the use of data-driven decision-making in creating class rosters. Despite the abundance of data available to practitioners, school leaders and staffs do not often systematically analyze available data to optimize classroom assignments or to ensure equitable access to effective teachers (Aaronson *et al.*, 2007; Burns and Mason, 2002; Kalogrides *et al.*, 2013; Kraemer *et al.*, 2011; Rothstein, 2009). This literature review suggests that data-driven decision-making could serve as a useful tool for improving classroom assignments.

School leaders could leverage data-driven decision-making in the rostering process to improve equitable access to effective teachers, for example. Existing rostering practices may yield inequity when higher performing students are systematically assigned to more effective teachers and lower performing students to less effective teachers (Aaronson *et al.*, 2007; Kalogrides *et al.*, 2013; Kelly, 2004; Koedel and Betts, 2009; Player, 2010). A contributing factor appears to be the greater ability of experienced and effective teachers to influence their principals' decisions on student assignments (Cohen-Vogel, 2011; Grissom *et al.*, 2015; Kraemer *et al.*, 2011). As a consequence, equitable opportunities to learn decrease and overall school achievement may be negatively impacted (Loeb *et al.*, 2012). Prior research indicates that principals have used measures of teacher effectiveness in rostering but only to remove ineffective teachers from tested grades (Chingos and West, 2011; Cohen-Vogel, 2011; Fuller and Ladd, 2013; Goldring *et al.*, 2015). School leaders could theoretically examine whether lower performing students are disproportionately assigned to less effective teachers. This review also highlights, however, the need to ensure that measures of teacher effectiveness are trustworthy indicators of teacher quality (Harris and Sass, 2014; MET Project, 2013; Whitehurst *et al.*, 2014).

School leaders and staffs could also use existing administrative data to examine the distributions of student achievement within and across classrooms. Even when school staffs reported efforts to balance classrooms on student characteristics, prior studies found that classrooms were unbalanced in terms of prior student achievement (Dieterle *et al.*, 2012; Kalogrides and Loeb, 2013). By gaining awareness of these tendencies, school leaders and staffs may consciously increase efforts to minimize negative peer effects that may occur when low-performing students are concentrated in any one class (Burke and Sass, 2013; Domina *et al.*, 2016; Imberman *et al.*, 2009; Vigdor and Nechyba, 2007). School leaders and staffs may also want to ensure that each student has peers in his/her class with similar achievement levels (Burke and Sass, 2013).

Another consideration is whether or not school leaders and staffs should strategically match specific students with teachers to optimize student learning. While research indicates that effective teachers were generally effective with all student subgroups (Lockwood and McCaffrey, 2009), prior studies indicate that strategic teacher-student pairings may improve student learning in some cases. Notably, consistently higher learning gains have been found in several studies when black students were assigned to a black teacher (Aaronson *et al.*, 2007; Dee, 2004; Egalite *et al.*, 2015; Gershenson *et al.*, 2017; Yarnell and Bohrnstedt, 2018). Along similar lines, ELs have been found to have greater learning gains when their teachers were fluent in Spanish and held bilingual certification, although prior research is limited on this point (Loeb *et al.*, 2014). Applying these findings to rostering, however, is not straightforward. For example, assigning black teachers a disproportionate number of black students could create negative effects if classrooms were then imbalanced in terms of other student attributes (Burke and Sass, 2013; Lavy and Schlosser, 2007; Van Ewijk and Sleegers, 2010).

While findings suggest that data-driven decision-making could improve the outcomes of the rostering process, systematic, data-driven rostering is not a commonly implemented practice. As such, any potential changes to routine rostering practices would need to be weighed against likely constraints concerning practicality of implementation (e.g. availability of data and capacity for analysis), acceptance by the school community (e.g. parents), and unintended consequences (e.g. teacher turnover). At this point in time, individual schools may simply lack the resources and tools for school staffs to systematically analyze existing data for the purpose of improving class rosters.

Data-driven rostering may also prove to be challenging in some contexts for other reasons. Prior research describes the conditions needed to facilitate data-driven decision-making in schools. School leaders need ongoing supports and incentives from districts (Anderson *et al.*, 2010; Hamilton *et al.*, 2009; Marsh, 2012; Snipes *et al.*, 2002). School staffs must be willing to engage in data-driven conversations in safe and trusting environments in which norms and beliefs can be examined (Coburn and Turner, 2011; Ikemoto and Marsh, 2007; Marsh, 2012), as underlying beliefs about class, race, and gender can pose a barrier to change (Valencia, 2010). Accordingly, district and school leader support is a necessary, though not entirely sufficient, ingredient for successful data-driven discussions in schools.

To the authors' knowledge, this study is the first exploration of data-driven decision-making in creating class rosters. This literature review identified several variables that could be used in data-driven rostering (e.g. teacher effectiveness, student achievement, student behavior, race and gender). A limitation of this review, however, is that it did not include all topics that may be relevant to the rostering process, such as class sizes. Data-driven rostering within a school requires further investigation.

This review outlined several ways in which school leaders and staffs could systematically analyze existing data to make improvements to class rosters, but the viability of these strategies for applied and expanded use in diverse schools remains untested. Future research could explore classroom groupings based on selected variables that optimized outcomes while limiting concomitant negative ripple effects (e.g. concentrations of chronically misbehaving students assigned to the same class). Additionally, researchers could simulate the gains or losses in student learning in certain rostering scenarios. Future research could also seek to understand how effects of specific rostering practices vary across different grade levels (e.g. elementary, middle, high) and school contexts.

Qualitative research is also needed to determine how school leaders could apply data-driven decision-making to the rostering process, and how school communities, including teachers and parents, react to rostering decisions made systematically. Qualitative research could inform the implementation barriers and unintended consequences of various rostering strategies, such as re-distributing effective teachers more equitably within schools, or using student characteristics more systematically to create classroom assignments. These lines of inquiry suggest the need for mixed-methods studies that examine both the implementation and impacts of different rostering strategies on various outcomes. Beyond student achievement, rostering effects on student behavior and social development, classroom and school climate, and teacher satisfaction should be explored.

In conclusion, our review of the literature suggests that data-driven rostering has the potential to yield more equitable outcomes than existing rostering practices. Yet the research base is limited, and there is not yet adequate research to support the exclusive use of data-driven rostering on the basis of quantitative data alone. More research on data-driven decision-making in creating class rosters is needed, along with associated professional development and guidance for school and district leaders.

References

Aaronson, D., Barrow, L. and Sander, W. (2007), "Teachers and student achievement in the Chicago public high schools", *Journal of Labor Economics*, Vol. 25 No. 1, pp. 95-135.

Anderson, S., Leithwood, K. and Strauss, T. (2010), "Leading data use in schools: organizational conditions and practices at the school and district levels", *Leadership and Policy in Schools*, Vol. 9, pp. 292-327.

Atteberry, A., Loeb, S. and Wyckoff, J. (2016), "High rates of within-school teacher reassignments and implications for student achievement", Working Paper [151], National Center for Analysis of Longitudinal Data in Education Research, Washington, DC.

Bacher-Hicks, A., Kane, T. and Staiger, D. (2014), "Validating teacher effect estimates using changes in teacher assignments in Los Angeles", Working Paper [20657], National Bureau of Economic Research, Cambridge, MA.

Berliner, D. (2009), *Poverty and Potential: Out-Of-School Factors and School Success*, Education Policy Research Unit, Arizona State University, Tempe, AZ.

Blazar, D. (2015), "Grade assignments and the teacher pipeline", *Educational Researcher*, Vol. 44 No. 4, pp. 213-227.

Bosworth, R. (2014), "Class size, class composition, and the distribution of student achievement", *Education Economics*, Vol. 22 No. 2, pp. 141-165.

Burke, M. and Sass, T. (2013), "Classroom peer effects and student achievement", *Journal of Labor Economics*, Vol. 31 No. 1, pp. 51-82.

Burns, R. and Mason, D. (2002), "Class composition and student achievement in elementary schools", *American Educational Research Journal*, Vol. 39 No. 1, pp. 207-233.

Cannata, M. (2010), "Understanding the teacher job search process: espoused preferences and preferences in use", *Teachers College Record*, Vol. 112 No. 12, pp. 2889-2934.

Cannata, M., Rubin, M., Goldring, E., Grissom, J.A., Neumerski, C.M., Drake, T.A. and Schuermann, P. (2017), "Using teacher effectiveness data for information-rich hiring", *Educational Administration Quarterly*, Vol. 53 No. 2, pp. 180-222.

Carey, N., Farris, E. and Carpenter, J. (1994), *Curricular Differentiation in Public High Schools*, National Center for Education Statistics, US Department of Education, Washington, DC.

Chingos, M. and West, M. (2011), "Promotion and reassignment in public school districts: how do schools respond to differences in teacher effectiveness?", *Economics of Education Review*, Vol. 30 No. 3, pp. 419-433.

Clotfelter, C., Ladd, H. and Vigdor, J. (2006), "Teacher-student matching and the assessment of teacher effectiveness", *Journal of Human Resources*, Vol. 41 No. 4, pp. 778-820.

Coburn, C.E. (2010), "Partnership for district reform: the challenges of evidence use in a major urban district", in Coburn, C.E. and Stein, M.K. (Eds), *Research and Practice in Education: Building Alliances, Bridging the Divide*, Rowman and Littlefield, New York, NY, pp. 167-182.

Coburn, C.E. and Turner, E.O. (2011), "Research on data use: a framework and analysis", *Measurement*, Vol. 9 No. 4, pp. 173-206.

Coburn, C.E. and Turner, E.O. (2012), "The practice of data use: an introduction", *American Journal of Education,* Vol. 118, pp. 99-111.

Cohen-Vogel, L. (2011), "Staffing to the test: are today's school personnel practices evidence based?", *Educational Evaluation and Policy Analysis*, Vol. 33 No. 4, pp. 483-505.

Condie, S., Lefgren, L. and Sims, D. (2014), "Teacher heterogeneity, value-added and education policy", *Economics of Education Review*, Vol. 40, pp. 76-92.

Dee, T.S. (2004), "Teachers, race, and student achievement in a randomized experiment", *The Review of Economics and Statistics*, Vol. 86 No. 1, pp. 195-210, available at: http://www.econis.eu/PPNSET?PPN=385839456.

Dieterle, S., Guarino, C., Reckase, M. and Wooldridge, J. (2012), "How do principals assign students to teachers? Finding evidence in administrative data and the implications for value-added", Discussion Paper 7112, The Institute for the Study of Labor, Bonn.

Domina, T., McEachin, A., Hanselman, P., Agarwal, P., Hwang, N. and Lewis, R. (2016), *Beyond Tracking and Detracking: The Dimensions of Organizational Differentiation in Schools*, RAND, Washington, DC.

Donaldson, M.L. and Papay, J.P. (2015), "An idea whose time had come: negotiating teacher evaluation reform in New Haven, Connecticut", *American Journal of Education*, Vol. 122, pp. 39-70.

Drake, T.A., Goldring, E., Grissom, J.A., Cannata, M., Neumerski, C.M., Rubin, M. and Schuermann, P. (2016), "Development or dismissal? Exploring principals' use of teacher effectiveness data", in Grissom, J.A. and Young, P. (Eds), *Improving Teacher Evaluation Systems: Making the Most of Multiple Measures*, Teachers College Press, New York, NY, pp. 116-130.

Egalite, A., Kisida, B. and Winters, M. (2015), "Representation in the classroom: the effect of own-race teachers on student achievement", *Economics of Education Review,* Vol. 45, pp. 44-52.

Engel, M., Jacob, B.A. and Curran, F.C. (2014), "New evidence on teacher labor supply", *American Educational Research Journal*, Vol. 51 No. 1, pp. 36-72.

Feng, L. (2010), "Hire today, gone tomorrow: new teacher classroom assignments and teacher mobility", *Education Finance and Policy*, Vol. 5 No. 3, pp. 278-316.

Figlio, D. (2007), "Boys named Sue: disruptive children and their peers", *Education Finance and Policy*, Vol. 2 No. 4, pp. 376-394.

Fox, L. (2016), "Playing to teachers' strengths: using multiple measures of teacher effectiveness to improve teacher assignments", *Education Finance and Policy*, Vol. 11 No. 1, pp. 70-96.

Fuller, S. and Ladd, H. (2013), "School-based accountability and the distribution of teacher quality across grades in elementary school", *Education Finance and Policy*, Vol. 8 No. 4, pp. 528-559.

Gershenson, S., Hart, C., Lindsay, C. and Papageorge, N. (2017), *The Long-Run Impacts of Same-Race Teachers: IZA Discussion Paper Series*, Institute of Labor Economics, Bonn, Germany.

Goe, L., Bell, C. and Little, O. (2008), *Approaches to Evaluating Teacher Effectiveness: A Research Synthesis*, National Comprehensive Center for Teacher Quality, Washington, DC, available at: http://www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf.

Goldhaber, D., Cowan, J. and Walch, J. (2013), "Is a good elementary teacher always good? Assessing teacher performance estimates across subjects", *Economics of Education Review*, Vol. 36, pp. 216-228.

Goldring, E., Grissom, J.A., Rubin, M., Neumerski, C.M., Cannata, M., Drake, T. and Schuermann, P. (2015), "Make room value added: principals' human capital decisions and the emergence of teacher observation data", *Educational Research*, Vol. 44 No. 2, pp. 96-104.

Green, B., Johnson, C. and Adams, A. (2006), "Writing narrative literature reviews for peer-reviewed journals: secrets of the trade", *Journal of Chiropractic Medicine*, Vol. 5 No. 3, pp. 101-117.

Grissom, J.A., Kalogrides, D. and Loeb, S. (2015), "The micropolitics of educational inequality: the case of teacher-student assignments", *Peabody Journal of Education*, Vol. 90, pp. 601-614.

Hamilton, L., Halverson, R., Jackson, S.S., Mandinach, E., Supovitz, J.A., Wayman, J.C., Pickens, C., Martin, E.S. and Steele, J.L. (2009), *Using Student Achievement Data to Support Instructional Decision Making*, United States Department of Education, Institute of Education Sciences, Washington, DC.

Hanushek, E., Kain, J., O'Brien, D. and Rivkin, S. (2005), "The market for teacher quality", Working Paper [11154], National Bureau of Economic Research, Cambridge, MA.

Hanushek, E., Kain, J. and Rivkin, S. (2002), "Inferring program effects for special populations: does special education raise achievement for students with disabilities?", *The Review of Economics and Statistics*, Vol. 84 No. 4, pp. 584-599.

Harris, D. (2006), *Lost Learning, Forgotten Promises*, Center for American Progress, Washington, DC.

Harris, D. (2009), "Would accountability based on teacher value added be smart policy? An examination of the statistical properties and policy alternatives", *Education Finance and Policy,* Vol. 4 No. 4, pp. 319-350.

Harris, D.N. and Herrington, C.D. (2015), "The use of teacher value-added measures in schools: new evidence, unanswered questions, and future prospects", *Educational Researcher,* Vol. 44 No. 2, pp. 71-76.

Harris, D.N. and Sass, T.R. (2014), "Skills, productivity and the evaluation of teacher performance", *Economics of Education Review*, Vol. 40, pp. 183-204.

Henderson, J. (2011), *Strategic Class Roster Creation in Elementary Schools: Indicative of Student Reading Growth or a Waste of Time? (Doctoral Dissertation)*, Wingate University, Matthews, NC.

Hill, A. (2014), "The costs of failure: negative externalities in high school course repetition", *Economics of Education Review*, Vol. 43, pp. 91-105.

Hopkins, G. (1999), *From the Principal Files: Making Class Lists Needn't Be a Nightmare!*, Education World, Colchester, CT, available at: http://www.educationworld.com/a_admin/admin/admin118.shtml.

Horoi, I. and Ost, B. (2015), "Disruptive peers and the estimation of teacher value added", *Economics of Education Review*, Vol. 49, pp. 180-192.

Hoxby, C. (2000), "Peer effects in the classroom: learning from gender and race variation", Working Paper [7867], National Bureau of Economic Research, Cambridge, MA.

Hoxby, C. and Weingarth, G. (2005), *Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects*, Harvard University, Cambridge, MA.

Ikemoto, G.S. and Marsh, J.A. (2007), "Cutting through the 'data-driven' mantra: different conceptions of data-driven decision making", *Evidence and Decision Making*, Vol. 106 No. 1, pp. 105-131.

Imberman, S., Kugler, A. and Sacerdote, B. (2009), "Katrina's children: evidence on the structure of peer effects from hurricane evacuees", Working Paper [15291], National Bureau of Economic Research, Cambridge, MA.

Ingersoll, R. (1999), "The problem of underqualified teachers in American secondary schools", *Educational Researcher,* Vol. 28 No. 2, pp. 26-37.

Ingram, D., Louis, K.S. and Schroeder, R.G. (2004), "Accountability policies and teacher decision making: barriers to the use of data to improve practice", *Teachers College Record*, Vol. 106, pp. 1258-1287.

Jackson, C. (2014), "Teacher quality at the high school level: the importance of accounting for tracks", *Journal of Labor Economics,* Vol. 32 No. 4, pp. 645-684.

Jiang, J.Y., Sporte, S.E. and Luppescu, S. (2015), "Teacher perspectives on evaluation reform: Chicago's REACH students", *Educational Researcher*, Vol. 44 No. 2, pp. 105-116.

Kahlenberg, R. (2004), *All Together Now: Creating Middle-Class Schools Through Public School Choice*. Brookings Institution Press, Washington, DC.

Kalogrides, D. and Loeb, S. (2013), "Different teachers, different peers: the magnitude of student sorting within schools", *Educational Researcher,* Vol. 42 No. 6, pp. 304-316.

Kalogrides, D., Loeb, S. and Béteille, T. (2013), "Systematic sorting: teacher characteristics and class assignments", *Sociology of Education*, Vol. 86 No. 2, pp. 103-123.

Kelly, S. (2004), "Are teachers tracked? On what basis and with what consequences", *Social Psychology of Education*, Vol. 7, pp. 55-72.

Kerr, K., Marsh, J., Ikemoto, G., Darilek, H. and Barney, H. (2006), "Strategies to promote data use for instructional improvement: actions, outcomes, and lessons from three urban districts", *American Journal of Education*, Vol. 112 No. 4, pp. 496-520.

Koedel, C. and Betts, J. (2009), "Value-added to what? How a ceiling in the testing instrument influences value-added estimation", Working Paper [14778], National Bureau of Economic Research, Cambridge, MA.

Kraemer, S., Worth, R. and Meyer, R. (2011), *Classroom Assignment Practices in Urban School Districts Using Teacher Level Value-Added Systems*, University of Wisconsin, Madison, WI.

Lavy, V. and Schlosser, A. (2007), *Mechanisms and Impacts of Gender Peer Effects at School*. Hebrew University of Jerusalem, Israel.

Lefgren, L. (2004), "Educational peer effects and the Chicago public schools", *Journal of Urban Economics,* Vol. 56, pp. 169-191.

Liu, K. (2010), *Peer Group Effects on Student Outcomes: Evidence from Randomized Lotteries (Doctoral Dissertation)*, Vanderbilt University, Nashville, TN.

Lockwood, J. and McCaffrey, D. (2009), "Exploring student-teacher interactions in longitudinal achievement data", *Education Finance and Policy*, Vol. 4 No. 4, pp. 439-467.

Loeb, S. and Candelaria, C.A. (2012), *How Stable Are Value-Added Estimates across Years, Subjects and Student Groups? (Knowledge Brief 3)*, Carnegie Foundation for the Advancement of Teaching, Stanford, CA.

Loeb, S., Kalogrides, D. and Béteille, T. (2012), "Effective schools: teacher hiring, assignment, development, and retention", *Education Finance and Policy*, Vol. 7 No. 3, pp. 269-304.

Loeb, S., Soland, J. and Fox, L. (2014), "Is a good teacher a good teacher for all? Comparing value-added of teachers with their English learners and non-English learners", *Educational Evaluation and Policy Analysis*, Vol. 36 No. 4, pp. 457-475.

Marsh, J. (2012), "Interventions promoting educators' use of data: research insights and gaps", *Teachers College Record*, Vol. 114, pp. 1-48.

Marsh, J., Hamilton, L. and Gill, B. (2008), "Assistance and accountability in externally managed schools: the case of edison schools, inc", *Peabody Journal of Education*, Vol. 83 No. 3, pp. 423-458.

Master, B. (2014), "Staffing for success: linking teacher evaluation and school personnel management in practice", *Educational Evaluation and Policy Analysis*, Vol. 36 No. 2, pp. 207-227.

MET Project (2013), *The Reliability of Classroom Observations by School Personnel*, Bill and Melinda Gates Foundation, Seattle, WA, available at: http://k12education.gatesfoundation.org/wp-content/uploads/2015/12/MET_Reliability-of-Classroom-Observations_Research-Paper.pdf.

Ost, B. (2014), "How do teachers improve? The relative importance of specific and general human capital", *American Economic Journal: Applied Economics*, Vol. 6 No. 2, pp. 127-151.

Papay, J. (2011), "Different tests, different answers: the stability of teacher value-added estimates across outcome measures", *American Educational Research Journal*, Vol. 48 No. 1, pp. 163-193.

Park, V., Daly, A.J. and Guerra, A.W. (2012), "Strategic framing: how leaders craft the meaning of data use for equity and learning", *Educational Policy*, Vol. 27 No. 4, pp. 645-675.

Player, D. (2010), "Nonmonetary compensation in the public teacher labor market", *Education Finance and Policy*, Vol. 5 No. 1, pp. 82-103.

Reardon, S. and Raudenbush, S. (2009), "Assumptions of value-added models for estimating school effects", *Education Finance and Policy,* Vol. 4 No. 4, pp. 492-519.

Rice, J. (2003), *Teacher Quality*, Economic Policy Institute, Washington, DC.

Rockoff, J.E., Staiger, D.O., Kane, T.J. and Taylor, E.S. (2012), "Information and employee evaluation: evidence from a randomized intervention in public schools", *The American Economic Review*, Vol. 102 No. 7, pp. 3184-3213.

Rothstein, J. (2009), "Student sorting and bias in value added estimation", Working Paper [14666], National Bureau of Economic Research, Cambridge, MA.

Rothstein, R. (2013), "Why children from lower socioeconomic classes, on average, have lower academic achievement than middle-class children", in Carter, P. and Welner, K. (Eds), *Closing the Opportunity Gap*, Oxford University Press, New York, NY, pp. 61-74.

Rusk, D. (2002), *Classmates Count: A Study of the Interrelationship between Socioeconomic Background and Standardized Test Scores of 4th Grade Pupils in the Madison-Dane County Public Schools*, Washington, DC, available at: http://www.schoolinfosystem.org/archives/Unifiedfinalreport.pdf.

Sacerdote, B., (2011), "Peer effects in education: how might they work, how big are they and how much do we know thus far?", in Hanushek, E., Machin, S. and Woessmann, L. (Eds), *Handbook of the Economics of Education*, North Holland, Amsterdam, Vol. 3, pp. 249-277.

Sanders, W. and Rivers, J. (1996), *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement*, Research Progress Report, University of Tennessee Value-Added Research and Assessment Center, Knoxville, TN.

Slavin, R.E. (1990), "Achievement effects of ability grouping in secondary schools: a best-evidence synthesis", *Review of Educational Research*, Vol. 60 No. 3, pp. 471-499.

Slavin, R.E., Cheung, A., Holmes, G., Madden, N.A. and Chamberlain, A. (2013), "Effects of a data-driven district reform model on state assessment outcomes", *American Educational Research Journal,* Vol. 50 No. 2, pp. 371-396.

Snipes, J., Doolittle, F. and Herlihy, C. (2002), *Foundations for Success: Case Studies of How Urban School Systems Improve Student Achievement*, MDRC for the Council of the Great City Schools, Washington, DC.

Spillane, J.P. (2012), "Data in practice: Conceptualizing the data-based decision-making phenomena", *American Journal of Education*, Vol. 118, pp. 113-141.

Steenbergen-Hu, S., Makel, M. and Olszewski-Kubilius, P. (2016), "What one hundred years of research says about the effects of ability grouping and acceleration on K-12 students' academic achievement: findings of two second-order meta-analyses", *Review of Educational Research*, Vol. 86 No. 4, pp. 849-899.

Valencia, R. (2010), *Dismantling Contemporary Deficit Thinking,* Routledge, New York, NY.

Van Ewijk, R. and Sleegers, P. (2010), "The effect of peer socioeconomic status on student achievement: a meta-analysis", *Educational Research Review*, Vol. 5 No. 2, pp. 134-150.

Vigdor, J. and Nechyba, T. (2007), "Peer effects in North Carolina public schools", in Woessmann, L. and Peterson, P.E. (Eds), *Schools and the Equal Opportunity Problem*, The MIT Press, Cambridge, MA, pp. 73-101.

Whitehurst, G., Chingos, M. and Lindquist, K. (2014), *Evaluating Teachers with Classroom Observations: Lessons Learned in Four Districts*, Brown Center on Education Policy at the Brookings Institution, Washington, DC.

Yarnell, L. and Bohrnstedt, G. (2018), "Student-teacher racial match and its association with Black student achievement: an exploration using multilevel structural equation modeling", *American Educational Research Journal*, Vol. 55 No. 2, pp. 287-324.

Corresponding author
Rebecca Wolf can be contacted at: betsywolf@jhu.edu

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**