

Using Non-Achievement Based Variables to Differentiate District Performance Grades

The COVID-19 pandemic disrupted many school accountability systems that rely on student-level achievement data. Many states have encountered uncertainty about how to meet federal accountability requirements without typical school data. Prior research provides an abundance of evidence that student achievement is correlated to students' social background, which raises concerns about the predictive bias of accountability systems. The focus of this quantitative study is to explore the predictive ability of non-achievement based variables (i.e., students' social background) on measures of school accountability in one Midwest state. Results suggest that social background and community demographic variables have a significant impact on measures of school accountability, and might be interpreted cautiously. Implications for policy and future research are discussed.

The role of accountability testing in education became widespread across the USA after the signing of the No Child Left Behind Act of 2001 (NCLB), and has continued throughout subsequent reauthorizations. States and government agencies have used student achievement data from large-scale standardized assessments to measure school and district performance. The COVID-19 pandemic, however, has disrupted the achievement-based accountability systems that were in place to some degree. “States face considerable uncertainty about how to meet federal and state accountability requirements for [the 2020-2021] school year and beyond” (Lake & Worthen, 2021, p. 1). This study examines one Midwest state’s report card accountability system. The report card accountability system assigns A-F letter grades as indicators of school and district performance. This study explores the predictive ability of non-achievement based variables (e.g., enrollment and median household income) on composite grades of school performance. The rationale for this study is two-fold. First, a statistical model based on social and community demographic (i.e., non-achievement) variables that can reliably predict measures of school accountability contributes evidence of the potential inequities permeating education systems. Thus, this research takes a strong access and equity perspective (National Council of Teachers of Mathematics, 2014). Second, exploring the predictive ability of social and community demographic variables encourages dialogue about whether current systems of accountability can operate without the achievement measures that have been impacted by the disruption from the COVID-19 pandemic. We seek to answer two research questions: (1) To what degree does a statistical model using non-achievement based variables reliably predict existing composite grades of school district performance? (2) How do non-achievement based variables differentiate the composite grades of school district performance?

Related Literature

Report Cards

Sixteen states, including the Midwest state that is the focus for this study, use A-F letter grades as a composite evaluation of school performance (Murray & Howe, 2017). The Midwest state’s administrative code defines the composite grade as the overall performance of a school or district (citation blinded). Table 1 documents the purposes of school and district report cards identified by the Midwest state’s department of education.

Student achievement data from end-of-grade (EOG) and end-of-course (EOC) assessments hold significant weight in the calculation of composite grades for schools and

districts. Figure 1 displays the six components that are used to determine composite grades. Student achievement, based on EOG and EOC standardized assessment results, is its own component of the report card, but interpretations of EOG and EOC assessment results are also used to measure other constructs such as *gap closing*. Gap closing measures the improvements in achievement by subgroups of students (citation blinded). Standardized assessment results are used in a multitude of ways, and are the dominant feature of school district performance in this model.

Previous Research of Student Achievement

Gaps in student achievement are intended to reflect differences in abilities, efforts, and/or the quality of education experiences provided to them. However, research has consistently indicated that differences in achievement are strongly correlated with students' social background (Broer et al., 2019; Caldas & Bankston, 1997; May, 2006). For example, socio-economic status (SES) is "the social standing or class of an individual or group" (APA, 2021). Caldas and Bankston's study of SES and academic achievement found attending school with classmates from high SES had a strong and significant correlation with academic achievement, regardless of one's own social background. That is, all students displayed greater academic achievement when attending a high SES school district (Caldas & Bankston, 1997). Broer and colleagues (2019) analyzed the relationship between SES and student achievement using twenty years of TIMSS data. A strong correlation between achievement scores and SES was reported. Previously, May (2006) reported a strong positive correlation between SES and fourth-grade reading achievement. Consequently, May (2006) questioned the validity of high stakes testing as a measure of student achievement.

Valid Interpretations and Uses

The *Standards* use the term *predictive bias* to describe differences "in the patterns of associations between test scores and other variables for different groups, bringing with it concerns about bias in the inferences drawn from the use of test scores" (AERA et al., 2014, p. 51). *Validity* can be defined as "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (AERA et al., 2014, p. 11). Validity is an attribute of the proposed interpretation and use of test scores, not the test itself (AERA et al., 2014). An argument-based approach to validation is necessary to support the interpretation and use of assessment results (AERA et al., 2014; Author, 2019; Author, in press; Cronbach, 1988; Kane, 2013). Additionally, "[Validation] needs to evaluate any potential sources of irrelevant variance in assessment scores and any potential negative consequences of score use" (Kane, 2020, p. 84). The degree to which a proposed interpretation and use is validated depends on the quantity and quality of evidence supporting the proposed interpretation and use (AERA et al., 2014; Author, 2019; Kane, 2013; 2020).

Studies that examine patterns of associations between test scores and other variables have implications for the validity of the intended interpretation and use of test scores (AERA et al., 2014). This study examines patterns of associations between composite grades of school district performance and non-achievement based variables (e.g., SES and race/ethnicity). A statistical model that reliably predicts composite grades of school performance may be used to encourage dialogue regarding the validity of the interpretation and use of test scores, Cronbach (1988) describes validation as a process of logical debate. This study contributes to the debate of

whether systems of accountability, such as A-F grades, support valid inferences about school quality and performance.

Methods

This quantitative study examined district-level data for all public school districts in Midwest state. Quantitative data consisted of non-achievement based data, such as social background and school district demographics, and composite grades of school district performance from 2018-2019. All data were publicly available through the state department of education. The variables of interest were selected based on prior literature and existing data.

The dependent categorical variable was composite grades of school district performance, which were collapsed into three categories. School districts earning a grade of A or B were combined into a “high-performing” category, and school districts earning a grade of D or F were combined into a “low-performing” category. This was done for two reasons. First, only 30 of the 600 school districts earned a grade of A, and only 4 of 600 school districts earned a grade of F. Second, a grade of A or B indicates a high performing district whereas a grade of D or F indicates a low performing district (blinded citation). Low performing districts are more likely to be subject to state-level interventions if achievement levels fail to improve. Data for 608 public school districts were retrieved. For a list of variables, see Appendix A. Eight districts were removed from the study due to missing data and/or districts identified as outliers due to small enrollment numbers.

A discriminant function analysis (DFA) was completed to determine if composite grades of school district performance could be determined by significant differences in non-achievement based variables. DFA was selected because it seeks to identify which combination of independent variables best predicts group membership. That is, DFA examines the relationship between a set of continuous independent variables and a categorical dependent variable (Tabachnick & Fidell, 2019). Standardized canonical coefficients were analyzed to address RQ2. Standardized canonical coefficients describe the correlation between independent variables and the discriminant functions. These coefficients represent the predictive contribution of a variable to a discriminant function (Tabachnick & Fidell, 2019).

Results

Preliminary analysis of descriptive statistics indicated SES and student attendance were related to composite grades of school performance. For example, the average real-estate property value per pupil was \$134,810 in school districts receiving a composite grade of D or F. Whereas the average property value per pupil was \$201,959 in school districts receiving a composite grade of A or B. Median household income also provided an interesting comparison between community SES and district performance. A community with a 2018 median household income of \$50,000 was three standard deviations above average-performing districts, but within one standard deviation of high-performing districts (i.e., districts receiving a composite grade of A or B). Table 2 displays descriptive statistics for variables related to SES and student attendance.

DFA revealed two discriminant functions. The first function explained 86.3% of the variance, canonical $R = 0.58$. The second function explained 13.7% of the variance, canonical $R = 0.18$. Thus, the two functions accounted for about 58% and 18% of the total relationship between independent variables and between composite grades. In combination, the discriminant functions significantly differentiated composite grades of school performance;

$\lambda = .$, $x () = .$, $p < .$. Removing the first function, the second function also significantly differentiated composite grades of school performance; $\lambda = .$, $x () = .$, $p < .$. In consideration of RQ1, the DFA produced a statistical model, combining non-achievement based variables, capable of predicting composite grades of school performance. Classification results reported 71.5% of original grouped cases were correctly classified. A casewise analysis of the classification results indicated the misclassification of a group was typically one level from the correct designation. That is, there were instances where a district earned a D/F composite grade, but the DFA predicted the district to earn a C. Analysis of standardized canonical coefficients indicated that attendance rate had a strong positive correlation with function 2, and districts' 'prepared for success' component grade had a moderate positive correlation with function 1. Table 3 displays the independent variables that were strongest in discriminant power. The larger the absolute value of the coefficient, the greater the variable contributes to discriminating groups (Tabachnick & Fidell, 2019).

Conclusions/Implications

The current study demonstrates that differences in non-achievement based variables, such as SES indicators, can predict composite grades of school performance. These results add to the findings of previous research regarding student achievement. That is, there exists an abundance of evidence that student achievement and social background variables are strongly correlated. This study extends beyond student achievement by demonstrating the predictive ability of social background variables on measures of school accountability. More specifically, this study presents evidence that it is possible to classify school performance without the use of student achievement data. This is important because current systems of accountability require the use of student-level achievement data. Perhaps the disruption to accountability systems caused by the COVID-19 pandemic provides an opportunity to consider how school districts are being held accountable.

Studies such as this are important in bringing attention back to the inequities permeating education. As stated by May (2006), "If we, as a nation, were to overtly acknowledge that wealth, or lack thereof plays a role in the success one is able to achieve, we would also have to acknowledge that some individuals are privileged by wealth and may even be bestowed with such at birth" (p. 52). One example of this are the discrepancies in median household income between low-, average-, and high-performing school districts. This research has implications for educational policy and future research. Results from this study suggest social background and community demographic variables influence measures of school accountability. Further research is warranted to explore the predictive bias of standardized assessment results, and subsequent accountability measures of school performance. The current study was limited by the availability of existing data. Future studies may consider the predictive ability of variables such as school climate and student engagement. From a policy perspective, this study may be used to encourage dialogue regarding the validity of school accountability systems, and how those systems may evolve to better represent the quality of education being provided to students.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*.

- American Psychological Association. (2021). *Socioeconomic status*. Retrieved from <http://www.apa.org/topics/socioeconomicstatus/>
- Author. (2019).
- Author. (in press).
- Broer, M., Bai, Y., & Fonseca, F. (2019). *Socioeconomic inequality and educational outcomes: Evidence from twenty years of TIMSS*. Springer.
- Blinded for Review [State Administrative Code]
- Blinded for Review [State Department of Education]
- Caldas, S. J., & Bankston, C. III. (1997). Effect of school population socioeconomic status on individual academic achievement. *The Journal of Educational Research*, 90(5), 269–277. <https://doi-org.ezproxy.bgsu.edu/10.1080/00220671.1997.10544583>
- Cronbach, L. J. (1988). *Five perspectives on validity argument*. In H. Wainer & H. Braun (Eds.), *Test Validity* (pp. 3-17). Erlbaum.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73. <https://doi.org/10.2307/23353796>
- Kane, M. T. (2020). Validity studies commentary. *Educational Assessment*, 25(1), 83-89. <https://doi.org/10.1080/10627197.2019.1702465>
- Lake, R., & Worthen, M. (2021) *State accountability systems in the COVID era and beyond*. Center on Reinventing Public Education. https://www.crpe.org/sites/default/files/v3_accountability_report_2021.pdf
- May, J. J. (2006). The role of money, race, and politics in the accountability challenge. *Journal of Urban Learning, Teaching, and Research*, 2, 46-55.
- Murray, K., & Howe, K. R. (2017). Neglecting democracy in education policy: A-F school report card accountability systems. *Education Policy Analysis Archives*, 25(109). <http://dx.doi.org/10.14507/epaa.25.3017>
- National Council of Teachers of Mathematics (2014). *Principles to actions: Ensuring mathematical success for all*.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, § 115 Stat. 1425. (2002). <http://www.ed.gov/legislation/ESEA02/>
- Tabachnick, B. G., & Fidell, L.S. (2019). *Using multivariate statistics*. Pearson.

Table 1

Purposes of the grade card accountability system (citation blinded)

Purpose	Description
Student growth and achievement	To provide communities a picture of school and district progress in raising student achievement and preparing students for the future.
Identify strengths and weaknesses	To provide educators, school administrators, and families information about the strengths and weaknesses of school performance.
Quality of education	To provide parents and schools an understanding about the quality of education being provided to students.

Figure 1

Components of school and district composite grades (citation blinded)

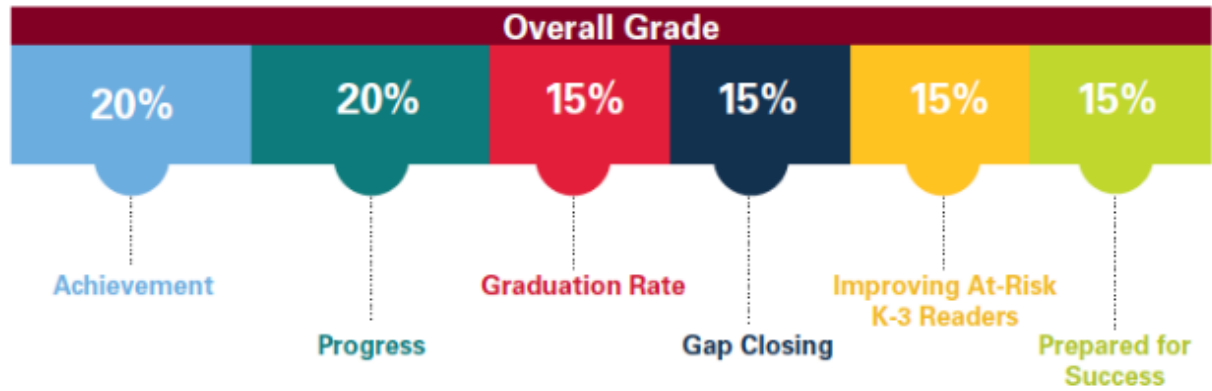


Table 2*Descriptive statistics of select variables*

Variable	Composite Grade	Mean	SD
Median household income	1 - D/F	\$31,398	4,196.8
	2 - C	\$35,262	4,616.2
	3 - A/B	\$43,262	10,617.8
Percent of students economically disadvantaged	1 - D/F	71.4%	24.46
	2 - C	44.65%	19.33
	3 - A/B	24.81%	14.95
Property value per pupil	1 - D/F	\$134,810	61,372
	2 - C	\$173,631	84,905
	3 - A/B	\$201,959	82,385
Chronic absenteeism	1 - D/F	20.67%	8.13
	2 - C	12.35%	5.08
	3 - A/B	7.57%	3.80
Student attendance	1 - D/F	92.7%	1.74
	2 - C	94.5%	1.10
	3 - A/B	95.5%	0.97

Table 3*Standardized canonical coefficients*

Function 1		Function 2	
Variable Description	Standardized Canonical Coefficient	Variable Description	Standardized Canonical Coefficient
Prepared for success component grade	0.362	Student attendance rate	0.704
Four-year graduation rate	0.359	Prepared for success component grade	0.645
Percent of students identified as economically disadvantaged	-0.285	Percent of students identified as Black	0.644
Average teacher salary	0.195	Chronic absenteeism	0.634

Appendix A

Description of Variables Used	
Dependent Variable	Composite grade of school district performance (low-, average-, high-performing)
Independent Variables	Median household income Number of students identified as economically disadvantaged Percent of students identified as economically disadvantaged 4-Year graduation rate 5-Year graduation rate Prepared for success percent score Student attendance rate Chronic absenteeism rate Percent of students residing within the district enrolled at the district Percent of students residing within the district open enrolled elsewhere Percent of students residing within the district attending a community school Number of school counselors per 1000 students Number of school psychologists per 1000 students Number of school social workers per 1000 students Number of general education teachers per 1000 students Number of intervention specialists per 1000 students Number of teacher aides per 1000 students Number of gifted intervention specialists per 1000 students Average property value per pupil Average teacher salary Average teacher experience Number of students identified as Black Percent of students identified as Black Number of students identified as Hispanic Percent of students identified as Hispanic Number of students identified as Multiracial Percent of students identified as Multiracial Number of students identified as White Percent of students identified as White Total Number of enrolled students