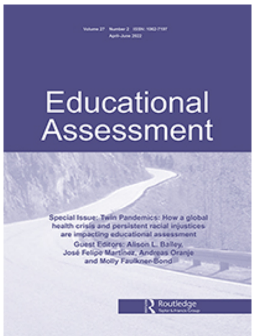

Educational Assessment

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/heda20

Allocating Resources for COVID-19 Recovery: A Comparison of Three Indicators of School Need

Jonathan D. Schweig, Andrew McEachin, Megan Kuhfeld, Lou Mariano & Melissa Diliberti

To cite this article: Jonathan D. Schweig, Andrew McEachin, Megan Kuhfeld, Lou Mariano & Melissa Diliberti (2022) Allocating Resources for COVID-19 Recovery: A Comparison of Three Indicators of School Need, Educational Assessment, 27:2, 152-169, DOI: 10.1080/10627197.2022.2087626

To link to this article: https://doi.org/10.1080/10627197.2022.2087626

EDUCATIONAL ASSESSMENT
2022, VOL. 27, NO. 2, 152–169
https://doi.org/10.1080/10627197.2022.2087626

Allocating Resources for COVID-19 Recovery: A Comparison of Three Indicators of School Need

Jonathan D. Schweig[a], Andrew McEachin[b], Megan Kuhfeld[b], Lou Mariano[a], and Melissa Diliberti[a]

[a]RAND Corporation, Santa Monica, California, USA; [b]NWEA, Portland, Oregon, USA

ABSTRACT

As students return to in-person instruction in the 2021–2022 school year, local education agencies (LEAs) must develop resource allocation strategies to support schools in need. Federal programs have provided resources to support restart and recovery. However, there is little consensus on how LEAs can target resources to support those schools most in need. This study investigates the relationship between three school need indicators (i.e., pre-COVID student performance and progress, school and community poverty, and pandemic vulnerability) and measures of student performance and progress throughout the pandemic to determine which indicators support valid school need inferences. We find that school poverty more strongly predicts performance and progress during the pandemic than pre-COVID academic measures. In elementary schools, we find that pandemic vulnerability independently predicts achievement even when conditioning on poverty and pre-pandemic achievement. Of the indicators of poverty we investigated, the percentage of free and reduced-price lunch-eligible students is the strongest predictor.

KEYWORDS

School accountability; assessment; COVID-19

The novel coronavirus disease 2019 (COVID-19) pandemic continues to have a profound impact on the lives of young people in the United States and across the globe. All 50 states closed schools to in-person instruction at some point in spring 2020 (Peele & Riser-Kositsky, 2020), and many students received at least some remote instruction through the end of the 2020–2021 school year (Institute of Education Sciences, n.d.). As a result, nearly 55 million school-aged children in the United States had reduced access to the social and academic supports typically provided by schools (García & Weiss, 2020).

The pandemic also threw a spotlight on the existing inequities in public education in the United States (Kaufman & Diliberti, 2021). Schools that struggled to provide high-quality learning under normal conditions found it even harder to teach effectively during the pandemic (García & Weiss, 2020). While evidence of how students overall are weathering the pandemic is still emerging, research shows students are learning at slower rates compared with prior years (Curriculum Associates, 2021; Lewis, Kuhfeld, Ruzek, & McEachin, 2021). Students in schools serving majority Black and Hispanic students, urban schools, and schools serving poorer communities were more likely to experience below-typical rates of growth during the first year of the pandemic compared with their White and higher-income peers (Curriculum Associates, 2021; Dorn, Hancock, Sarakatsannis, & Viruleg, 2020).

COVID-19 has disproportionately impacted poorer communities and Black, Hispanic, and Indigenous communities (Dorn et al., 2020). Inequities in the social determinants of health, including health-care access, put these communities at increased risk of contracting and dying from COVID-19 (Centers for Disease Control and Prevention [CDC], 2021). Black, Hispanic, and Indigenous

CONTACT Jonathan D. Schweig jschweig@rand.org RAND Corporation, 1776 Main Street, Santa Monica, CA 90407
Supplemental data for this article can be accessed online at https://doi.org/10.1080/10627197.2022.2087626

households were more likely to struggle with food insufficiency during the pandemic, and renters of color faced the greatest hardships in terms of their ability to pay rent or mortgage (Center on Budget and Policy Priorities, 2020). Exposure to these kinds of stressors can have adverse developmental impacts on children, both in terms of academic achievement and in terms of social and emotional well-being (e.g., Goodman, Miller, & West-Olatunji, 2012).

Additionally, internet access and device availability presented a significant challenge for schools using fully remote or hybrid approaches, especially at the beginning of the pandemic: nationally, only 30% of teachers in high-poverty schools reported that all or nearly all their students had internet access at home, compared with 83% of teachers in low-poverty schools (Stelitano et al., 2020). Fully remote instruction was more common in urban schools (Schwartz et al., 2021) and schools with higher shares of Black and Hispanic students, students who experience homelessness, and students that are free or reduced-price lunch (FRPL) eligible (Parolin & Lee, 2021; Zamarro & Camp, 2021). The concentration of remote learning among historically marginalized student populations is concerning because preliminary reports (e.g., Kaufman & Diliberti, 2021) suggest remote learning throughout the pandemic was generally less rigorous than in-person schooling.

As educators reckon with the pandemic's impact on schooling and schools return to in-person instruction for the 2021–2022 school year (Belsha & Barnum, 2021), local education agencies (LEAs) must develop resource allocation strategies to provide supports to students and schools in need (National Academy of Education, 2021). Despite early predictions that the pandemic would force budget reductions in many districts (Burnette, 2020), many places were able to avoid severe spending cuts (e.g., Lieberman, 2021). One reason why school budgets held up better than expected is that the federal government allocated billions of dollars to state education agencies (SEAs) across three stimulus packages: the Coronavirus Aid, Relief, and Economic Security (CARES) Act; the Coronavirus Response and Relief Supplemental Appropriations Act (CRRSA); and the American Rescue Plan Act (ARP), which made the largest contribution to the Elementary and Secondary School Emergency Relief (ESSER) Fund. All three stimulus packages allocated funds to SEAs based on the Title I, Part A (Title I) formula of the Elementary and Secondary Education Act, meaning that the country's least advantaged school districts, on average, received greater amounts of this funding (Barnum & Belsha, 2021). Funds are intended to support operations, offer mental health services to students, and address disrupted learning in 2021–2022 and beyond.[1]

While allocating funds to those students and schools that were hit hardest by the pandemic is a priority, the task of identifying which schools and students would benefit most from additional funding has largely been delegated to LEAs. Accordingly, LEAs have a lot of autonomy to determine how these resources should be allocated across schools in 2021–2022 and subsequent school years. Nearly 80% of LEA funding (about $88 billion) provided through the ARP can be spent at the discretion of LEAs (Policy Innovators in Education Network, 2021). However, LEAs had to develop resource allocation strategies for ARP funds quickly. For the 2021–2022 school year, reopening plans had to be published within 30 days of receipt of funds, annual school year 2021–2022 budgets needed local approval by the end of June 2021, and LEAs had to submit their ARP plans to states within 90 days of receipt of funds (Roza et al., 2021). Such an accelerated timeline left LEAs little time to consult with stakeholders and make data-informed spending decisions. LEAs will follow a similar process for the 2022–2023, 2023–2024, and 2024–2025 school years.

LEAs might balance two dimensions when considering how to best allocate funds. The first is guided by a federal accountability policy (i.e., the Every Student Succeeds Act [ESSA]) that dictates how states and districts must identify low-performing schools and allocate supportive resources. While states and districts have flexibility in school accountability system design, school need is typically identified by a composite indicator of academic achievement and growth. The U.S.

[1]CRRSA funds must be allocated and spent by Sept 2023, and CARES and ARP funds must be allocated and spent by September 2024.

Department of Education (ED) has paused formal school accountability at least through 2020–2021 with schools carrying forward their pre-pandemic performance designations. We provide more context on this federal guidance in the background section.

The second dimension includes the three channels through which the pandemic affected formal schooling: (1) the physical closure of schools and associated disruptions to typical teaching practices; (2) the impact of closing large parts of the economy to ensure the public's health and safety, which disproportionately affected poorer communities; and (3) the direct impact on community health and well-being. LEAs have multiple data sources to help proxy these channels that might guide allocation decisions, including (1) estimates of school performance and rate of progress using pre-pandemic assessment data; (2) estimates of school and community poverty including school Title I status or the proportion of FRPL-eligible students; and (3) estimates of pandemic vulnerability (e.g., Snyder & Parks, 2020).

A fourth indicator of schools' need for additional support could be obtained from estimates of school performance and rate of progress using spring assessment data from the 2020–2021 school year – a use that is explicitly supported by federal guidance (U.S. Department of Education, 2021c). However, emerging evidence suggests that spring 2021 test participation was low and uneven. Of the 30 states that have released test participation data thus far, less than half have had participation rates of 90% or more, and some states saw participation rates as low as 10% (Center for Reinventing Public Education, 2021). Several states have reported that Black and economically disadvantaged students were less likely to participate in spring 2021 assessments than their White and more advantaged peers (Barnum, 2021; EPIC, 2021). If spring 2021 assessment data are not available for many students in an LEA, this compromises the ability of officials to track and monitor growth (Fazlul, Koedel, Parsons, & Qian, 2021b). Given continued disruptions to schooling and assessment, many LEAs will not have access to robust assessment data that can be used to measure performance and progress and inform resource allocation decisions. As of the writing of this article, the assessment policy for 2021–2022 remains unclear, further compromising the availability of assessment data for LEAs.

Each of these indicators (i.e., pre-COVID performance and progress, estimates of school and community poverty, and pandemic vulnerability) has advantages and limitations, and important questions remain concerning which indicator (or combination of indicators) is most predictive of school need in the pandemic era. There are many ways in which states and districts may define school need. As we discuss in more detail shortly, we follow the spirit of federal accountability policy to define need as schools with lower rates of school performance and rate of progress (i.e., a weighted average of school-aggregated test scores and measures of achievement growth). Which indicators LEAs use to proxy this definition of need, and how they use them, will further depend on how districts want to target resources.

Current study

This study investigates the empirical relationships between each of these indicators with measures of student performance and progress throughout the pandemic as evidence of whether these indicators allow for valid inferences about school need. We then investigate how combinations of indicators can potentially be used to inform resource allocation decisions. Using a sample of nearly 1.7 million students in grades 3 to 8 who took NWEA's MAP® Growth™ assessments in the 2017-2018, 2018-2019, 2019–2020 and 2020–2021 school years (about 7% of the approximately 22 million U.S. public school students in grades 3 to 8 according to the U.S. Department of Education, National Center for Education Statistics, 2021b), we address two specific research questions:

1. To what extent do indicators of pre-COVID performance and progress, school and community poverty, and pandemic vulnerability allow for valid inferences about school need? Specifically, how predictive are each of these indicators of school performance and progress during the pandemic?

2. How can these indicators be combined to improve predictions of school need?

To preview our results, we find that school poverty and estimates of pandemic vulnerability more strongly predict school-level academic achievement during the pandemic than pre-COVID academic measures. In elementary schools only, we find that indicators of pandemic vulnerability independently predict school performance and progress even when conditioning on school poverty and pre-pandemic school achievement. Of the three specific indicators of poverty that we investigate, the percentage of FRPL-eligible students is the strongest predictor of school performance and progress during the pandemic.

In the following sections, we provide an overview of prior research and outline the contributions of this paper. We provide details on the three need indicators, provide illustrative examples of how LEAs are using these indicators, and outline the strengths and weaknesses of each indicator as suggested by the literature. We then describe our conceptual framework, including how we situate our investigations as part of a validity argument about each indicator as a measure of school need. We proceed with a discussion of the data and methods used to address the research questions. We then present the results of our analyses and conclude with a discussion of the implications of our findings for making accurate resource allocation decisions.

Background

In a typical (pre-pandemic) school year, assessment data played a critical role in identifying low-performing schools (or school need) and determining how resources are allocated to schools. For example, as a component of the school accountability systems mandated by ESSA, states must designate Comprehensive Support and Improvement (CSI) schools and Targeted Support and Improvement (TSI) schools based on the indicators in the state's accountability system. CSI and TSI statuses are used to identify struggling schools, and LEAs are required to develop plans to support school improvement and address resource inequities for CSI and TSI designated schools. In many states, a weighted composite of school performance (or achievement levels) and rate of progress (or achievement growth) is the central component of the state's accountability system (Lyons, D'Brot, & Landl, 2017; Portz & Beauchamp, 2020).

Physical school closures across the U.S. in spring 2020 threw annual state assessment programs into disarray. Recognizing the potential effects that COVID-19 would have on student learning and wellbeing, U.S. ED granted waivers in spring 2020 to all 50 states to allow SEAs and LEAs to bypass annual ESSA-mandated assessments (Gewertz, 2020). While U.S. ED required all states to conduct assessments in some form in 2020–2021, the Department once again offered accountability waivers and granted flexibility in terms of when and how these assessments were administered and in what form (U.S. Department of Education, n.d.).

In this context where LEAs have atypical, if any, assessment data but must identify schools in need to receive additional resources, we identified three plausible indicators that LEAs might consider to make valid school-need predictions and move forward with resource allocation decisions for the upcoming school years. As mentioned earlier, there are several ways to define school need. However, given that states and districts have worked under a federal accountability framework since the implementation of the No Child Left Behind Act in 2002 and will likely once again have to identify low-performing schools using an accountability framework in the 2021–2022 school year and beyond, we use a definition of school need that closely aligns with pre-pandemic federal policy guidance. Below, we provide a brief description of each indicator to highlight important features and to provide a justification for our inclusion of each of these indicators in this study. We also provide illustrative examples of existing LEA policies that make use of such indicators and summarize some of the key underlying assumptions of each indicator.

School performance and rate of progress using pre-pandemic assessment data

One way for LEAs to determine which of their schools have the greatest need of support is to use pre-pandemic assessment data to estimate school performance and rate of progress. Such a method has been used by SEAs and LEAs in compliance with U.S. ED waiver policy in both the 2019–2020 and 2020–2021 school years. As a condition for those waivers, state officials had to attest that schools identified as CSI or TSI in the 2019–2020 school year would maintain that identification status in the 2020–2021 school year and continue to receive supports from the LEA consistent with the school's support and improvement plan (U.S. ED, n.d.).

Some SEAs also suspended policies for determining performance indicators for the 2020–2021 school year and are instead carrying forward these indicators from the pre-pandemic 2018–2019 school year. For example, in Ohio, House Bill 164 uses a Community School of Quality designation to allocate additional funding to schools, but to receive the designation (and accompanying funding), schools need to have met academic requirements from 2017 to 2018 and 2018 to 2019 (Ohio House Bill 164, Ohio Department of Education, 2021). In Colorado, House Bill 21–1161 suspends the determination of performance indicators for schools until July 2022, meaning that formal determinations of school performance and progress will not be updated until fall 2022 (Colorado House Bill 1161, Colorado Department of Education, 2021). Similarly, in Arkansas, House Bill 1151 suspends the public-school rating system for the 2020–2021 school year, carrying forward school ratings from 2018 to 2019 (Arkansas House Bill 1151).

Carrying forward pre-pandemic data about school performance to inform resource allocation for the 2021–2022 school year is consistent with a "hold harmless" philosophy: specifically, as it is unlikely any schools improved their performance during the pandemic, it is reasonable to assume schools that were identified for assistance before the pandemic would remain in need for 2019–2020 and 2020–2021. Under the assumption that schools that needed improvement before the pandemic are still struggling, carrying forward pre-pandemic data about school performance and progress may help LEAs ensure that resources are efficiently allocated.

School and community poverty

As described above, CARES, CRRSA, and ARP allocate funding to states based on the Title I formula, which sends extra money to schools serving children from low-income families (Snyder, Dinkes, Sonnenberg, & Cornman, 2019). LEAs could likewise use school Title I status or other indicators of student poverty, such as FRPL eligibility, to inform resource allocation decisions. There are many good reasons to use indicators of student poverty for such purposes. First, it is well established that socioeconomic status is one of the strongest predictors of academic achievement (e.g., Reardon, 2013). We also know that the needs of high-poverty schools and districts differ from the needs of low-poverty schools and districts (e.g., EdTrust, n.d.). Second, Title I status and FRPL eligibility data are readily available to LEAs and SEAs to inform decision-making. Some school districts already use school poverty data to inform resource allocation decisions. For example, district leaders in Boston Public Schools rely on an Opportunity Index, which incorporates more fine-grained information about neighborhood poverty and student socioeconomic status, to direct resources to the highest-need schools (Boston Public Schools, 2019).

Despite widespread use of these indicators, research has shown that both Title I status and FRPL eligibility tend to overstate school poverty. Schoolwide Title I status may overstate poverty because only 40% of a school's students must be FRPL-eligible to qualify (Dynarski & Kainz, 2015). FRPL eligibility may also overstate school poverty, as eligibility rates exceed expectations based on income (Fazlul, Koedel, & Parsons, 2021a; Harwell & LeBeau, 2010). Though FRPL eligibility is a coarse measure, research does show that FRPL eligibility captures aspects of educational disadvantage not captured by other poverty measures (Domina et al., 2018).

COVID-19 vulnerability

Assessment data and data on school and community poverty are widely used in education research and practice. These indicators, however, do not directly account for the extent to which specific communities were impacted by COVID-19. We posit that community-level pandemic vulnerability influenced students' opportunity to learn (i.e., that schools in locations hit hard by the pandemic are in greater need of resources for recovery). There are many factors beyond race and socioeconomic status that may make communities vulnerable to COVID-19 (Nayak et al., 2020), including factors like obesity rates, air quality, hospital bed availability, and the extent to which communities were successful at social distancing (Snyder & Parks, 2020). Emerging research shows that while race and poverty explain a significant amount of variation in COVID-19 vulnerability from county to county, there are persistent differences beyond those explained by these factors alone (Adhikari et al., 2020; Ramchand et al., 2019; Snyder & Parks, 2020).

Because pandemic vulnerability is driven by a combination of ecological, social, health, and economic variables, another approach to determining which schools or school communities may be in greater need of resources to support pandemic recovery is to look at a social vulnerability index that uses multiple variables to directly quantify community vulnerability to COVID-19. For example, Snyder and Parks (2020) use a weighted composite of 19 variables including temperature, air quality, education, hospital beds, race, obesity, and poverty as a vulnerability index. The National Institute of Environmental Health Sciences (NIEHS) publishes a COVID-19 Pandemic Vulnerability Index (PVI) that monitors disease trajectories and communicates local vulnerability (Marvel et al., 2021). The PVI is a composite index from 12 indicators including current infection rates, population concentration, existing policies for interventions, and health and environmental vulnerabilities (including race and poverty).

Conceptual framework

The process of validation entails accumulating evidence supporting the interpretation of a measure for specific uses, including making inferences about students, teachers, and schools (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014; Baker & Linn, 2002; Kane, 2006), and should attend to the expected benefits and unintended consequences that may arise for each proposed use (Messick, 1989). Each of the three indicators described above can be used to make inferences about school need and thus to inform resource allocation decisions, including staff hiring, adopting educational programs, or investing in instructional materials, among others. Accordingly, we consider our examinations of these indicators as providing validity evidence supporting their use in informing these kinds of decisions. We argue that appraising the strength of the relationships among these need indicators and a measure of school performance and progress provides key evidence on whether, in the absence of a typical testing regime, these indicators can provide useful information to support resource allocation decisions.

Specifically, this study investigates the validity of these three school need indicators in terms of criterion-related evidence. Our criterion measure of school need is a composite of student performance and progress during the pandemic, derived from winter 2020 and winter 2021 interim assessment data. School need during the pandemic is complex, and stakeholders may differ both in terms of which aspects of school need to prioritize and how to operationalize the target variables. Ideally, we would be able to construct additional criterion measures that reflect a wide range of valued student outcomes (e.g., students' physical, social, and emotional well-being). However, the pandemic not only disrupted traditional schooling structures but also disrupted the collection of student data, so we do not have access to such measures. Additionally, there are different ways to operationalize school need, including, for example, changes in growth trajectories relative to pre-pandemic or secular trends. Our criterion variable selection was strongly informed by our belief that LEAs are interested in using a definition of school need that is consistent with how need was commonly operationalized in the pre-pandemic period.

Evidence regarding the relationships among the need indicators and our criterion measures is an important source of validity information because the strength of these relationships provides a way to understand the extent to which inferences about school need based on each indicator (or combination of indicators) align with inferences about school need based on the criterion measure directly. When the association between a need indicator and the criterion measure is strong, inferences about school need will be similar. When the association between a need indicator and the criterion measure is weak, inferences may start to diverge (Kane & Staiger, 2002; Mihaly, McCaffrey, Staiger, & Lockwood, 2013).

Valid predictions of school need are critical to support LEAs as they work toward restart and recovery, so that school need indicators provide the best support possible for making resource allocation strategies (D'Brot, Landl, Domaleski, & Brandt, 2020). This is particularly important in the context of COVID-19 because there is such wide variation in how the pandemic has impacted communities. If there are schools that have seen dramatic shifts in their ability to effectively meet the academic needs of their students as they work to overcome pandemic-induced stressors, a successful resource allocation strategy should accurately identify this need and ensure those schools receive the support they need.

Methods

Analytic sample

Data for this study come from NWEA's anonymized longitudinal student achievement database. We use data for 1,654,340 third- to eighth-grade students in 2,966 public schools within eight U.S. states. While NWEA administers the MAP® Growth™ assessments in 47 states (and the District of Columbia), we restricted our sample to include only states that had a sufficiently large number of schools that participated in MAP Growth assessments in four school years (2017–2018, 2018–2019, 2019–2020, and 2020–2021). We made this restriction because we estimate school performance and progress (see Analytic Methods below) separately by state, subject, and grade level, and high-quality estimates require a reasonable number of participating students and schools. We began state selection by ranking states by coverage level (i.e., the proportion of their schools that participated in MAP Growth assessment) and focused on selecting states with high coverage. We then selected states that would maximize the sample heterogeneity in terms of school characteristics (e.g., urbanicity, student FRPL eligibility, and majority White vs. Black vs. Hispanic), and geography (Northeast vs. Midwest vs. West vs. South). Finally, we tried to include states that had been impacted differently by COVID-19 (using observed case rates and hospitalizations as a proxy for impact).

The NWEA database includes student-level demographic information, including race/ethnicity, gender, and age at assessment. Information about student-level socioeconomic status is not available. Table 1 provides descriptive statistics for the student sample by grade. Overall, the sample is 51% male, 54% White, 13% Black, 4% Asian, and 16% Hispanic, with similar demographic patterns observed across grade levels. As a point of comparison, the projected percentage distribution of students enrolled in public schools in the 2020–2021 school year in the eight states in our sample is 48% White, 18% Black, 6% Asian, and 23% Hispanic (U.S. ED, 2021a). While our analytic sample is generally representative of the population of public school students in selected states, it has a slight overrepresentation of White students and a slight underrepresentation of Hispanic students compared with the state populations.

Information about school-level characteristics, including the proportion of FRPL-eligible students, was obtained from the 2019–2020 Common Core of Data (CCD) files (U.S. ED, 2021b).[2] Around 75% of the public schools in the 2020–2021 school year in the eight states in our sample are Title I schools, and around 50% of the students are FRPL-eligible. Our analytic

[2]Some schools and districts use the Community Eligibility Provision (CEP) and consider all of their students to be FRPL-eligible. All schools using the CEP have 100% FRPL eligibility in our sample.

Table 1. Student sample demographic characteristics by grade.

	Male	White	Black	Hispanic	Asian	Native American	Hawaiian	Multi-race	Other
Overall	51.0%	53.9%	12.8%	15.9%	3.5%	0.9%	0.2%	3.8%	9.0%
Grade 2	50.8%	56.0%	11.8%	13.6%	3.5%	0.9%	0.2%	3.9%	10.1%
Grade 3	51.1%	53.2%	12.4%	16.5%	3.5%	0.9%	0.2%	4.0%	9.3%
Grade 4	50.9%	53.2%	13.4%	16.2%	3.5%	1.0%	0.2%	3.9%	8.6%
Grade 5	50.8%	54.7%	12.8%	15.7%	3.5%	0.9%	0.2%	3.6%	8.6%
Grade 6	51.0%	51.6%	13.2%	18.2%	3.9%	0.9%	0.2%	3.7%	8.3%
Grade 7	51.0%	53.2%	13.5%	17.5%	3.4%	1.1%	0.2%	3.4%	7.7%
Grade 8	51.2%	52.7%	14.3%	16.7%	3.3%	0.9%	0.2%	3.4%	8.5%

Note. N = 1,654,340.

sample is generally representative of the public school population in our selected states. However, the NWEA sample has a slight overrepresentation of Title I schools and a slight underrepresentation of FRPL-eligible students compared with the state populations (Table 2).

Typical testing policies use a "spring to spring" timeline where students are tested at the end of each school year, and progress is defined as students' growth between spring tests. To maintain a year-long period, we use a "winter to winter" timeline given spring 2020 scores are unavailable.[3] To determine whether the students who participated in MAP testing in both winter 2020 and winter 2021 were similar to those in the sample in pre-COVID assessment periods, we analyzed the student- and school-level match rates in both pre-pandemic and pandemic periods. We calculated the percentage of students in the same schools that could be matched to valid test scores in adjacent school years. We found student-level match rates to be generally similar across periods, suggesting that, while test participation was somewhat lower in winter 2021 than in winter 2020, students did not drop out of our analytic sample in the COVID era in ways that might bias our results.[4] For example, among third graders with a test observation in spring 2017, we matched 55% of them with a 4th-grade test observation in the same school in spring 2018. Between spring 2018 and spring 2019, the match rate among 3rd graders was 59%, and the match rate between winter 2020 and winter 2021 was 50%.

Table 2. School-level descriptive statistics for analytic sample.

	Elementary				Middle			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Z score math 2019	0.04	0.47	−3.15	1.82	0.03	0.52	−2.81	2.84
Z score reading 2019	0.05	0.41	−2.99	1.81	0.00	0.47	−3.47	1.90
Rate of progress math 2019	0.00	0.12	−1.85	1.64	0.03	0.16	−1.08	1.56
Rate of progress reading 2019	0.00	0.10	−2.67	1.07	0.02	0.18	−1.76	2.13
Achievement summary 2019	0.00	0.72	−8.43	3.23	0.00	0.69	−4.17	3.81
Title I	0.77	0.42	0.00	1.00	0.75	0.43	0.00	1.00
Percent of students eligible for free or reduced-price lunch (FRPL)	0.46	0.24	0.01	1.00	0.53	0.27	0.02	1.00
Poverty composite	0.00	2.27	−5.50	9.21	0.00	2.19	−5.20	8.52
Pandemic Vulnerability Index	0.49	0.06	0.25	0.60	0.51	0.06	0.28	0.60
Z score math 2021	−0.21	0.54	−3.31	2.28	−0.11	0.57	−2.92	2.29
Z score reading 2021	−0.06	0.51	−2.85	1.64	−0.04	0.54	−2.71	1.59
Rate of progress math 2021	−0.01	0.16	−1.45	1.65	−0.01	0.13	−1.22	1.36
Rate of progress VAM reading 2021	−0.01	0.13	−1.62	1.35	0.01	0.09	−0.82	1.31
Achievement summary 2021	0.00	0.73	−6.42	2.96	0.00	0.65	−3.86	4.29

Note. N for elementary = 2,081. N for middle = 1,483.

[3]We acknowledge the choice of test timeline can influence the identification of low and/or high-performing schools (e.g., McEachin & Atteberry, 2017). However, this is true not only during the pandemic but during prior school years. A winter to winter timeline is the closest timeline available to the typical spring to spring timeline.

[4]While other analyses of MAP data (e.g., Lewis et al., 2021) have found evidence of systematic differences in student participation based on student demographic characteristics that differed relative to prior years, these analyses apply different sample inclusion criteria and are not directly comparable to the sample in the current analysis. Analyses of MAP Growth data using similar sample inclusion criteria (Schweig et al., 2022) have found similar attrition patterns to those described in this study.

Similar patterns were observed among students in other grade levels and also when we considered match rates at the school level versus the student level. For a more in-depth comparison of MAP Growth participation rates before and during the COVID-19 pandemic, including by student demographics and community risk factors, see Schweig, Kuhfeld, Diliberti, McEachin, and Mariano (2022) and Kuhfeld, Ruzek, Lewis, and McEachin (2021).

Instruments

The primary instruments for this analysis are NWEA's MAP Growth assessments in mathematics and reading. MAP Growth is a computer adaptive test that measures achievement even for students above or below grade level and is vertically scaled to allow for the estimation of gains within and across grade levels. MAP Growth assessments are typically administered three times a year (fall, winter, and spring) and are aligned to state content standards (NWEA, n.d.). Test scores are reported on the RIT (Rasch unIT) scale, which is a linear transformation of the logit-scale units from the Rasch item response theory model. School districts use MAP Growth assessments to monitor students' reading and mathematics growth throughout the school year. MAP Growth scores strongly correlate with test scores on state summative end-of-year assessment (correlations ranging from .80 to .86 across most states, see NWEA, 2019).

Indicators of school need

In this section, we provide details about each of the school-need indicators we use to predict our criterion measure of school need, including information about how the indicators were constructed. Descriptive information on these indicators for our analytic sample of schools is available in Table 2.

Pre-pandemic school performance and rate of progress

We calculated two measures of school achievement (performance and progress) separately for math and reading using spring 2018 and spring 2019 achievement data. We standardized the test scores using NWEA's national norms (see Thum & Kuhfeld, 2020) and created separate performance measures for math and reading that were simple school-level averages across the 2018 and 2019 achievement data.

Our progress measure used a value-added model, similar to models used in ESSA-compliant accountability frameworks (e.g., SAS, 2017), often termed a "two-step" model or average residual model (e.g., Ehlert, Koedel, Parsons, & Podgursky, 2016; Koedel, Mihaly, & Rockoff, 2015). The first step of the model is an OLS regression of students' math or reading spring scores on their prior math and reading scores from the prior spring, as well as school-level averages of math and reading for their school from the prior year:

$$Y_{ist} = \beta_0 + \beta_1 Y_{is,t-1} + \beta_2 \tilde{Y}_{is,t-1} + \beta_3 \bar{Y}_{is,t-1} + \beta_4 \bar{\tilde{Y}}_{is,t-1} + \varepsilon_{ist} \,. \tag{1}$$

For student i in school s at time t we regressed a student's math or reading score, Y_{ist}, on her same-subject, $Y_{is,t-1}$, and off-subject, $\tilde{Y}_{is,t-1}$, score from the prior year as well as her school's average lagged same-subject, $\bar{Y}_{is,t-1}$, and off-subject, $\bar{\tilde{Y}}_{is,t-1}$, scores. We estimated this model separately by grade and state (e.g., State A third grade).

The second step took the student-level estimated residuals from (1), $\hat{\varepsilon}_{ist}$, pooled the data across grades (3 to 5 for elementary schools and 6 to 8 for middle schools), and used OLS to estimate a simple regression of $\hat{\varepsilon}_{ist}$ on a vector of dummy variables for all schools in a state:

$$\hat{\varepsilon}_{ist} = \theta' D_s + \omega_{ist} \,, \tag{2}$$

where D_s is a fully saturated vector of school-specific dummy variables that took a value of 1 if a student attended a given school and a value of zero otherwise; ω_{ist} is an idiosyncratic error-term. We interpreted the vector of coefficients, $\hat{\theta}$, as our measure of school progress. These coefficients captured the average student achievement growth for a given school conditional on students' prior achievement. We created separate averages for elementary and middle grades. If a school included both, it had separate measures of performance and progress for each.

Finally, we used a standard empirical Bayes shrinkage technique (e.g., Aaronson, Barrow, & Sander, 2007; as discussed in Koedel et al., 2015) to account for noise due to sampling variation: $\hat{\theta}_s^{EB} = \delta_s\hat{\theta}_s + (1 - \delta_s)\bar{\theta},$ where δ_s is a measure of the signal to noise ratio and $\bar{\theta}$ is the mean progress measure across schools.[5]

We did not control for students' race/ethnicity in this model to align our practices with the accountability models allowed under federal guidance. Ehlert et al. (2016) demonstrate that lagged school-level averages of achievement are sufficient statistics for school-level demographics in a school value-added model.

With each of the four measures in hand (performance and progress for both math and reading), we created a summary measure of school-level achievement similar to the indices used in accountability policies (as described above). We did this by standardizing each of the four measures to put them on the same metric and then created a simple average.

School and community poverty

We used three measures of school and community poverty. Title I status was obtained from the 2019–2020 CCD files. We define Title I status as any school eligible for participation in Title I programs. We consider both schools eligible for targeted assistance programs and schoolwide Title I programs to be "Title I" schools. FRPL eligibility describes the percentage of students in each school that are eligible to receive free or subsidized meals as part of the National School Lunch Program (a federal program administered by the U.S. Department of Agriculture designed to provide healthy meals to children from low-income families). Data on school-level FRPL eligibility were also obtained from the CCD.

Finally, we used a set of district characteristics that were calculated and reported by the Stanford Education Data Archive (SEDA) Version 4.0. District characteristics from SEDA provide information both on district resources and the characteristics of the community residing within the school district geographic boundaries (for details, see Reardon et al., 2021). District-level characteristics pulled from the SEDA database include median family income, percentage of adults with a bachelor's degree or higher, poverty rate, unemployment rate, percentage of households receiving Supplemental Nutrition Assistance Program (SNAP) benefits, and percentage of single-mother households. We used a Principal Components Analysis to reduce the dimensions of the SEDA characteristics into a summary variable. The first component explained 64.26% (60.08%) of the variation and had an eigenvalue of 5.14 (4.81) with weights essentially equal across the variables for elementary (middle) schools.

Vulnerability to COVID-19

We used the COVID-19 Pandemic Vulnerability Index (PVI) as a county-level measure of community vulnerability. The PVI is published daily by the NIEHS and is comprised of four domains: infection rate, population concentration, intervention measures (e.g., social distancing and testing), and health and environment. The fourth domain – health and environment – contains information on the percentage of the population that identifies as Black or American Indian (for more information

[5]Here $\delta_s = \frac{\hat{\sigma}_\theta^2}{\hat{\sigma}_\theta^2 + \hat{\Delta}_s}$ is the ratio of the variance in school-specific progress over the sum of this variance and a school's specific error variance. We estimate the former,$\hat{\sigma}_\theta^2$, by taking the variance in the school progress measures, $var\left(\hat{\theta}_s\right)$, and subtracting out the mean of the squared standard errors of the school progress measures: $\hat{\sigma}_\theta^2 = var\left(\hat{\theta}_s\right) - mean\left(\hat{\Delta}_s\right)$. We estimate the former as simply the square of the school's standard error: $\hat{\Delta}_s = se\left(\hat{\theta}_s\right)^2$..

about this index, see the NIEHS website). We pulled the PVI each day from February 28, 2020 (the first day the PVI was published) to April 7, 2021 (our cutoff date for beginning analysis) and took the average.

Criterion measure: pandemic performance and rate of progress

As described above, we used a composite measure of performance and progress that closely aligns with federal accountability guidance as our school needs a criterion measure. To calculate pandemic performance and rate of progress, we follow a process similar to the pre-pandemic measures of school-level achievement with two key differences. First, the pandemic school-level rate of progress is calculated for only one school year, using winter 2021 as students' outcomes controlling for winter 2020 scores as students' lagged achievement. Second, given that students largely were not assessed in spring 2020 – and because we did not yet have spring 2021 test scores at the time of this analysis – we use a winter-to-winter timeline instead. Specifically, we replaced students' spring scores in Equation 1 with students' winter 2021 scores and students' prior year spring scores with students' winter 2020 scores. The rest of the procedures (i.e., averaging of residuals and empirical Bayes shrinkage) remain the same.

Analytic methods

We estimate simple school-level OLS regressions of our criterion measure on our indicators of school need: (1) pre-COVID school-level progress and performance; (2) school and community poverty (separate models for Title I status, FRPL eligibility, and a community poverty composite); and (3) a county-level indicator of COVID-19 vulnerability. All models include state fixed-effects and use Hubert-White robust standard errors.

Our regressions provide empirical evidence of the strength of the associations among these indicators of school need and our criterion measure. We interpret stronger relationships between a need indicator and the criterion measure as providing stronger support for the use of a specific indicator in predicting school need. To gauge the strength of a relationship, we examine the magnitude of both the unstandardized and standardized regression coefficients, as well as the coefficient of determination (R^2), which quantifies the proportion of variance in the criterion measure that is explained by our predictors. The model coefficients show the extent to which each need indicator is predictive of school performance and progress during the pandemic.

Our school-level OLS regressions do not account for school-level race/ethnicity for a few reasons. First, the channels through which LEAs are likely able to allocate resources include measures of academic performance and/or poverty (e.g., FRPL eligibility and/or Title I status). Second, our PVI measure includes race/ethnicity, and both achievement and school-level poverty are strongly correlated with race/ethnicity. Finally, while there are good arguments to use these additional educational resources to help ameliorate racial injustices, our data are not well equipped for this type of analysis. LEAs may want to use local ethnicity as a fourth dimension through which/ethnicity as a fourth dimension through which to allocate resources.,[6,7]

Results

We first present results comparing the extent to which the three indicators of school need are predictive of our criterion measure. We then present results about the combination (or combinations) of indicators that are most useful for supporting resource allocation decisions. As a sensitivity analysis,

[6]As a sensitivity analysis, we conducted all regressions using student race and ethnicity variables. Race/ethnicity was not a statistically significant predictor of our criterion measures. Results are available upon request.

[7]A related question is the extent to which our results vary across states or other contexts. This is an important line of inquiry. Given the limited number of states in our sample, we are unable to adequately examine state-level heterogeneity (e.g., random coefficient model) and leave this analysis to future research.

Table 3. Summary of regressions using need indicators to predict criterion.

Predictor	Coefficient	Standard error	Standardized coefficient	R^2
Elementary schools				
Achievement[a]	0.398	0.033	0.396***	0.157
FRPL[a]	−1.563	0.060	−0.524***	0.251
Title I[b]	−0.474	0.040	−0.275***	0.070
Poverty[c]	−0.137	0.007	−0.430***	0.169
PVI[a]	−2.571	0.291	−0.196***	0.038
Middle schools				
Achievement[d]	0.323	0.032	0.343***	0.117
FRPL[d]	−0.863	0.062	−0.362***	0.116
Title I[e]	−0.310	0.042	−0.205***	0.037
Poverty[f]	−0.085	0.008	−0.290***	0.074
PVI[d]	−0.484	0.273	−0.047	0.002

Note. FRPL: free or reduced price lunch; PVI: Pandemic Vulnerability Index. [a]$N = 2081$. [b]$N = 2053$. [c]$N = 2019$. [d]$N = 1483$. [e]$N = 1460$. [f]$N = 1430$. ***$p < .001$.

we also conducted these regressions separately on the constituent components of our criterion measure and by subject (e.g., separately on performance and progress in reading and math). The results are similar for these additional analyses and largely reinforce our findings. Thus, for the sake of brevity, we present results for our achievement composite criterion in this section and provide additional results in an online appendix.

Results from elementary schools (Table 3) indicate that each of the need indicators had a statistically significant relationship with our criterion measure ($p < .001$). The highest predictive association was between FRPL eligibility and the criterion measure, and the lowest predictive association was between PVI and the criterion measure. Results from middle schools (Table 3) also indicate that the need indicators have statistically significant relationships with the criterion measure ($p < .001$) though PVI is not a significant predictor of the criterion.

Table 4 provides results of the multiple linear regressions using pre-pandemic achievement, poverty indicators, and pandemic vulnerability to simultaneously predict our criterion measures. Model 1 uses pre-pandemic achievement in combination with FRPL eligibility and PVI. Model 2 uses pre-pandemic achievement in combination with Title I status and PVI. Model 3 uses pre-pandemic achievement in combination with our poverty composite and PVI. In elementary schools, the regression model that used FRPL eligibility in combination with PVI and pre-pandemic achievement (Model 1) explained 29.3% of the variance in the criterion measure. This is the largest percentage

Table 4. Comparison of regressions using combined need indicators to predict criterion.

Predictor	Model 1	Model 2	Model 3
Elementary schools			
Achievement	0.192*** (0.025)	0.329*** (0.029)	0.257*** (0.029)
FRPL	−1.172*** (0.077)		
Title I		−0.334*** (0.040)	
Poverty			−0.088*** (0.009)
PVI	−1.886*** (0.295)	−2.995*** (0.323)	−2.148*** (0.318)
R^2	0.293	0.224	0.239
N	2081	2053	2019
Middle schools			
Achievement	0.243*** (0.0306)	0.297*** (0.0318)	0.256*** (0.0300)
FRPL	−0.695*** (0.0756)		
Title I		−0.222*** (0.0408)	
Poverty			−0.0637*** (0.009)
PVI	0.474 (0.344)	−0.807* (0.313)	0.187 (0.352)
R^2	0.179	0.137	0.143
N	1483	1460	1430

Note. Standard errors appear in parentheses. FRPL: free or reduced price lunch; PVI: Pandemic Vulnerability Index. ***$p < .001$; *$p<.05$.

of variance explained by any of the models (22.4% for Model 2 and 23.9% for Model 3) and larger than any of the univariate models we considered. In middle schools, the story is largely the same. Model 1 explained the largest percentage of variance in the criterion measure (17.9% compared to 13.7% for Model 2 and 14.3% for Model 3).

Summary

Using MAP Growth test scores collected prior to and during the COVID-19 pandemic, we investigated the degree to which three potential indicators of school need predicted student academic performance and progress throughout the pandemic. We then explored which combination(s) of indicators could be combined to improve these predictions. Although our need indicators explain less of the variance in middle school performance and progress than elementary, our results show that 1) each of our need indicators independently predicts our criterion measures; 2) FRPL eligibility is the strongest such predictor; and 3) the most accurate predictions of progress and performance during the pandemic are made by using a combination of pre-pandemic academic indicators, FRPL eligibility, and pandemic vulnerability. For elementary schools, PVI independently predicted our criterion variable even when controlling for pre-pandemic school performance and progress and school-level poverty.

When we considered all three need indicators simultaneously, FRPL eligibility was a stronger predictor of the criterion measure than either PVI or pre-pandemic performance and progress. Our results demonstrate that, even when pre-pandemic academics and community pandemic vulnerability are accounted for, the marginal relationship between FRPL eligibility and the criterion measures shows that FRPL eligibility represents an important aspect of school need.

Two of these results are particularly intriguing. First, it is noteworthy that FRPL eligibility emerges as the strongest predictor of our criterion measures even though, in the education literature, prior achievement is often the strongest predictor of current achievement. We expect different measures of the same construct to have higher correlations than similar measures of different constructs (e.g., Campbell & Fiske, 1959). In the current study, we might anticipate that our pre-pandemic measures of performance and progress would be the strongest predictors of our criterion measures. In other words, we might expect that schools where students had been making academic progress prior to the pandemic would tend to be the schools where students had made progress during the pandemic. In fact, in analyses available upon request, we find a cross-year correlation for our criterion variable using pre-pandemic data of approximately .75, significantly higher than in our analysis. However, that is not the case for our criterion variable across the pandemic. Instead, we find that one measure of school poverty – FRPL eligibility – best predicts our criterion measures. This could mean that the pandemic disrupted the strong performance and progress for some schools with higher shares of students in poverty, effectively strengthening the correlation between poverty and COVID-era achievement and weakening the correlation between pre-pandemic achievement and COVID-era achievement.

Second, given how the pandemic has disproportionately hurt traditionally underserved communities and students, it is not surprising that school poverty is a strong predictor of our criterion variable. However, it is somewhat unexpected that of the three measures of school and community poverty that we investigated, it was FRPL eligibility that was the strongest predictor of our criterion measure, given that recent research has problematized the use of FRPL eligibility as a poverty measure (e.g., Domina et al., 2018; Fazlul et al., 2021a; Michelmore & Dynarski, 2017). Fazlul and colleagues (2021) note that FRPL eligibility rates often exceed expectations based on income thresholds for participation and recommend using measures of poverty based on family income, similar to the kinds of information included in our poverty composite. One potential explanation for this phenomenon is that FRPL eligibility may capture aspects of educational disadvantage that are not captured by information about household income. Such an interpretation is supported by Domina and colleagues (2018), who found that, conditional on a student's family income, FRPL indicators have a statistically and practically significant negative relationship with achievement. We interpret this as evidence that FRPL eligibility is sensitive to the challenges faced by students throughout the pandemic.

Given that a combination of pre-pandemic academic indicators, FRPL eligibility, and pandemic vulnerability (PVI) provides relatively accurate predictions of progress and performance during the pandemic, the results of this study suggest that this combination of indicators can be useful in informing resource allocation decisions and determining school need. We caution that though this combination of indicators was the most predictive of the set that we compared, the majority of the variance in the criterion measure was unexplained by our predictors, which suggests that these indicators should not form the sole basis of any school need decisions. Consistent with calls by assessment experts, LEAs should rely on multiple sources of data – and in particular on sources that are highly sensitive to the local context – to inform strategic decision-making during the pandemic (D'Brot et al., 2020).

Limitations

We acknowledge several limitations of this study. First, while we selected states that had high levels of coverage of MAP Growth testing, the schools in our sample are not necessarily representative of all public schools in the state. Districts choose to offer MAP Growth for a variety of reasons; that selection process may result in both observed and unobserved differences between schools selecting into testing and those that did not. Second, the MAP Growth tests can differ in meaningful ways from the end-of-year summative tests that are used in accountability systems, including the content, stakes, and timing. While MAP Growth scores have been shown to strongly relate to end-of-year assessment scores in multiple states (NWEA, 2021), it is possible that results would have been different if we had access to summative assessment scores. Third, some students and schools move in and out of the sample across the years. While we performed preliminary analyses to establish that student and school pandemic-period mobility was similar to pre-pandemic trends, we are not able to fully account for differences in the test-taking population from school to school or district to district, as well as differences in test administration mode (e.g., remote or in-person), in assessment length, and the test administration window. Additionally, the fact that we were only able to match around half of our students across test administrations may compromise the generalizability of our growth measures to the entire school. While growth model-based estimates of school progress may not be sensitive to this kind of attrition (Fazlul et al., 2021b), low test participation threatens the validity of school-level inferences. Fourth, academic performance and progress are only one potential criterion measure that may be of importance. Other criteria, such as student social and emotional well-being, may be important as well, and one limitation of our study is that the MAP Growth data do not include measures of students' non-test outcomes. Our understanding of the accuracy of the school need indicators depends strongly on how need is defined and how the target criterion is operationalized. Fifth, there are some plausible alternative explanations for our findings. Our measures of school need are observed at different levels (e.g., PVI is a county-level measure and FRPL is a school-level measure) and contain a mixture of both continuous and dichotomous variables. We cannot rule out the possibility that some results are due to differences in how these variables are defined, and future work should interrogate this possibility to the extent possible.

Finally, we note that there are many important sources of validity evidence that are not considered in this investigation, including evidence based on content, response processes, and internal structure. Decisions to use any of the resource allocation strategies described above should be informed by evidence from all these sources. However, such an investigation is beyond the scope of the current study, and we do not have the data needed to investigate these issues thoroughly. Even given these limitations, we believe that our results have important implications for the allocation of resources during the 2021–2022 school year and beyond.

Conclusion

The COVID-19 pandemic has reshaped public and private life, including primary and secondary schooling. We are just starting to unpack the negative effects of the pandemic on students' opportunities to learn mental and physical well-being and other aspects of their lives. At the same time, schools entered the 2021–2022 school year with unique financial resources at their disposal to mitigate some of the negative consequences of the pandemic.

The pandemic affected formal schooling through (1) the physical closure of schools and associated disruption to typical teaching practices; (2) the impact of closing large parts of the economy to ensure the public's health and safety; and (3) the direct impact on community health and well-being. It will be important for LEAs to account for all three of these factors as they use additional resources to implement COVID-19 recovery programs, policies, and interventions.

The results of this study have implications for how LEAs can use readily available administrative data to inform resource allocation decisions. School and community contexts play an important role in predicting COVID-era school performance and progress, highlighting the importance of using a broad set of indicators to design and implement COVID-19 recovery programs, policies, and interventions in the intervening school years. In particular, information about both school-level poverty (specifically FRPL eligibility) and COVID-19 vulnerability allow LEAs not only capture a good deal of the variation in school-level performance and progress but also the combination of these measures is likely to perform better as a predictor of school need than any single indicator we investigated.

Finally, although our criterion measures are focused on achievement and do not directly include non-academic outcomes, we suspect that both FRPL eligibility and COVID-19 vulnerability are strongly correlated with these other outcomes that may be of policy interest or local importance. In this manner, using information about FRPL eligibility and COVID-19 vulnerability to inform resource allocation strategies would likely allocate resources toward school and students most impacted by the pandemic in other ways.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The research reported here was sponsored by the Institute of Education Sciences, U.S. Department of Education, through grant [R305U200006]. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95–135. doi:10.1086/508733

Adhikari, S., Pantaleo, N. P., Feldman, J. M., Ogedegbe, O., Thorpe, L., & Troxel, A. B. (2020). Assessment of community-level disparities in coronavirus disease 2019 (COVID-19) infections and deaths in large US metropolitan areas. *JAMA Network Open*, *3*(7). https://doi.org/10.1001/jamanetworkopen.2020.16938

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

Baker, E. L., & Linn, R. L. (2002). *Validity issues for accountability systems*. Los Angeles, California: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University of California, Los Angeles. Retrieved from https://cresst.org/wp-content/uploads/TR585.pdf

Barnum, M. (2021, February 24). This year's state test results will be tough to make sense of, experts warn. *Chalkbea*t. Retrieved from https://www.chalkbeat.org

Barnum, M., & Belsha, K. (2021, March 25). Unprecedented federal funding is on the way. High-poverty schools are starting to reckon with how to spend it. *Chalkbeat*. Retrieved from https://www.chalkbeat.org

Belsha, K., & Barnum, M. (2021, May 3). 'I expect it.' Cardona says every school needs in-person instruction this fall. *Chalkbeat*. Retrieved from https://www.chalkbeat.org

Boston Public Schools. (2019). *The opportunity index*. Retrieved from https://www.bostonpublicschools.org/cms/lib/MA01906464/Centricity/Domain/2301/Opportunity%20Index%20overview%20fact%20sheet%202.1.19.pdf

Burnette, D., II (2020, May 8). Devastated budgets and widening inequities: How the coronavirus collapse will impact schools. *Education Week*. Retrieved from https://www.edweek.org

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81. doi:10.1037/h0046016

Center on Budget and Policy Priorities. (2020, August 21). *Tracking the COVID-19 recession's effects on food, housing, and employment hardships*. Retrieved from https://www.jstor.org/stable/resrep25613

Center on Reinventing Public Education. (2021). *State Guidance on School Reopenings amid the Pandemic*. Webpage.

Centers for Disease Control and Prevention. (2021, April 19). *Health equity considerations and racial and ethnic minority groups*. Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html

Colorado Department of Education. (2021, May 6). *Accountability pause for the 2020-2021 and 2021-2022 school years*. Retrieved from https://www.cde.state.co.us/accountability/accountabilitypause

Curriculum Associates. (2021). *What we've learned about unfinished learning: Insights from midyear diagnostic assessments*. Retrieved from https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-understanding-student-needs-paper-winterresults-2021.pdf

D'Brot, J., Landl, E., Domaleski, C., & Brandt, C. (2020). *Restart and recovery considering the outlook for school accountability: State guidance for making annual accountability determinations in school year 2020–2021 and beyond*. Council of Chief State School Officers. Retrieved from https://www.nciea.org/sites/default/files/inline-files/CCSSO_RR_Outlook_Accountability-v4.pdf

Domina, T., Pharris-Ciurej, N., Penner, A. M., Penner, E. K., Brummet, Q., Porter, S. R., & Sanabria, T. (2018). Is free and reduced-price lunch a valid measure of educational disadvantage? *Educational Researcher*, *47*(9), 539–555. doi:10.3102/0013189X18797609

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020, June 1). *COVID-19 and student learning in the United States: The hurt could last a lifetime*. McKinsey & Company. Retrieved from https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-student-learning-in-the-united-states-the-hurt-could-last-a-lifetime

Dynarski, M., & Kainz, K. (2015). *Why federal spending on disadvantaged students (Title I) doesn't work*. Brookings. Retrieved from https://www.brookings.edu/wp-content/uploads/2016/07/Download-the-paper-2.pdf

EdTrust. (n.d.) *Resource allocation reviews: A critical step to school improvement*. Retrieved from https://edtrust.org/wp-content/uploads/2014/09/ESSA_FactSheet__Resource-Allocation-Reviews_Hyperlinks-UPDATED.pdf

Education Policy Innovation Collaborative. (2021). *Michigan's 2020–21 Benchmark Assessments*. East Lansing, Mich.

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2016). Selecting growth measures for use in school evaluation systems: Should proportionality matter? *Educational Policy*, *30*(3), 465–500. doi:10.1177/0895904814557593

Fazlul, I., Koedel, C., & Parsons, E. (2021a). *Free and reduced-price meal eligibility does not measure student poverty* (CALDER Working Paper No. 252-0521). Center for Analysis of Longitudinal Data in Education Research. Retrieved from https://caldercenter.org/sites/default/files/CALDER%20WP%20252-0521.pdf

Fazlul, I., Koedel, C., Parsons, E., & Qian, C. (2021b). Estimating test-score growth for schools and districts with a gap year in the data. *AERA Open*, *7*(1), 1–14. doi:10.1177/23328584211041346

García, E., & Weiss, E. (2020). *COVID-19 and student performance, equity, and U.S. education policy: Lessons from pre-pandemic research to inform relief, recovery, and rebuilding*. Economic Policy Institute. Retrieved from https://files.eric.ed.gov/fulltext/ED610971.pdf

Gewertz, C. (2020, April 2). It's official: All states have been excused from statewide testing this year. *Education Week*. Retrieved from https://www.edweek.org

Goodman, R. D., Miller, M. D., & West-Olatunji, C. A. (2012). Traumatic stress, socioeconomic status, and academic achievement among primary school students. *Psychological Trauma: Theory, Research, Practice, and Policy*, *4*(3), 252–259. doi:10.1037/a0024912

Harwell, M. R., & LeBeau, B. (2010). Student eligibility for a free lunch as an SES measure in education research. *Educational Researcher*, *39*(2), 120–131. doi:10.3102/0013189X10362578

Institute of Education Sciences. (n.d.) *Monthly school survey dashboard*. Retrieved from https://ies.ed.gov/schoolsurvey/

Kane, T. J., & Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives*, *16*(4), 91–114. doi:10.1257/089533002320950993

Kane, M. T. (2006). Validation. In R. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). New York, NY: Praeger.

Kaufman, J., & Diliberti, M. K. (2021). *Divergent and inequitable teaching and learning pathways during (and perhaps beyond) the pandemic: Key findings from the American Educator Panels Spring 2021 COVID-19 Surveys*. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA168-6.html

Koedel, C., Mihaly, K., & Rockoff, J. (2015). Value-added modeling: A review. *Economics of Education Review*, *47*(10), 180–195. doi:10.1016/j.econedurev.2015.01.006
Kuhfeld, M., Ruzek, E., Lewis, K., & McEachin, A. (2021). *Technical appendix for learning during COVID-19: Reading and math achievement in the 2020-21 school year*. Portland, OR: NWEA.
Lewis, K., Kuhfeld, M., Ruzek, E., & McEachin, A. (2021). *Learning during COVID-19: Reading and math achievement in the 2020-21 school year*. NWEA. Retrieved from https://www.nwea.org/content/uploads/2021/07/Learning-during-COVID-19-Reading-and-math-achievement-in-the-2020-2021-school-year.research-brief.pdf
Lieberman, M. (2021, March 19). School budgets: Why they're not as bad as predicted. *Education Week*. Retrieved from https://www.edweek.org
Lyons, S., D'Brot, J., & Landl, E. (2017). *State systems of identification and support under ESSA: A focus on designing and revising systems of school identification*. Council of Chief State School Officers. Retrieved from https://ccsso.org/sites/default/files/2017-12/State%20Systems%20of%20ID%20and%20Support%20-%20Designing%20and%20Revising%20Systems_0.pdf
Marvel, S. W., House, J. S., Wheeler, M., Song, K., Zhou, Y. H., Wright, F. A., ... Reif, D. M. (2021). The COVID-19 pandemic vulnerability index (PVI) dashboard: Monitoring county-level vulnerability using visualization, statistical modeling, and machine learning. *Environmental Health Perspectives*, *129*(1), 017701. doi:10.1289/EHP8690
McEachin, A., & Atteberry, A. (2017). The impact of summer learning loss on measures of school performance. *Education Finance and Policy*, *12*(4), 468–491. doi:10.1162/edfp_a_00213
Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York, NY: American Council of Education and Macmillan.
Michelmore, K., & Dynarski, S. (2017). The gap within the gap: Using longitudinal data to understand income differences in educational outcomes. *AERA Open*, *3*(1), 233285841769295. doi:10.1177/2332858417692958
Mihaly, K., McCaffrey, D. F., Staiger, D., & Lockwood, J. R. (2013). *A composite estimator of effective teaching*. Seattle, WA: Bill & Melinda Gates Foundation. https://www.rand.org/pubs/external_publications/EP50155.html
National Academy of Education. (2021, March 4). Educational assessments in the COVID-19 era and beyond. [*PowerPoint slides*]. Retrieved from https://naeducation.org/wp-content/uploads/2021/03/Educational-Assessments-in-the-COVID-19-Era-and-Beyond_NAEd-Forum-Combined-Slides.pdf
Nayak, A., Islam, S. J., Mehta, A., Ko, Y., Patel, S. A., Goyal, A., , and Quyyumi, A. A. (2020). *Impact of social vulnerability on COVID-19 incidence and outcomes in the United States*. medRxiv. doi:10.1101/2020.04.10.20060962
NWEA. (2019). MAP Growth technical report. Portland, OR: Author. Retrieved from https://www.nwea.org/content/uploads/2021/11/MAP-Growth-Technical-Report-2019_NWEA.pdf
NWEA. (2021). Predicting performance on the smarter balanced summative assessments based on NWEA MAP growth scores. Retrieved from https://www.nwea.org/content/uploads/2021/03/SBAC-MAP-Growth-Linking-Study-Report-2021-03-30.pdf
NWEA. (n.d.) *MAP growth instructional areas*. Retrieved November 18, 2021, from https://cdn.nwea.org/state-information/index.html#/
Ohio Department of Education. (2021). *Quality community schools support criteria and frequently asked questions*. Retrieved from https://education.ohio.gov/Topics/Community-Schools/Quality-Community-School-Support-Fund/Quality-Community-Schools-Support-Criteria-Frequen
Parolin, Z., & Lee, E. K. (2021). Large socio-economic, geographic and demographic disparities exist in exposure to school closures. *Nature Human Behavior*, *5*(4), 522–528. doi:10.1038/s41562-021-01087-8
Peele, H., & Riser-Kositsky, M. (2020, September 16). Map: Coronavirus and school closures in 2019–2020. *Education Week*. Retrieved from https://www.edweek.org
Policy Innovators in Education Network. (2021). *Distribution of education funds in the American Rescue Plan*. Retrieved from https://pie-network.org/wp-content/uploads/2021/03/ARP-Education-Funding-Overview.pdf
Portz, J., & Beauchamp, N. (2020). Educational accountability and state ESSA plans. *Educational Policy*, *36*(3), 0895904820917364.
Ramchand, R., Seelam, R., Parks, V., Ghosh-Dastidar, B., Lee, M. R., & Finucane, M. (2019). Exposure to the Deepwater Horizon oil spill, associated resource loss, and long-term mental and behavioral outcomes. *Disaster Medicine and Public Health Preparedness*, *13*(5–6), 889–897. doi:10.1017/dmp.2019.3
Reardon, S. F. (2013). The widening income achievement gap. *Educational Leadership*, *70*(8), 10–16.
Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., Jang, H., & Chavez, B. (2021). *Stanford education data archive: Technical documentation (version 4.0)*. Center for Education Policy Analysis, Stanford University. Retrieved from http://purl.stanford.edu/db586ns4974
Regards student religious expression, Ohio House B. 164. (2020).
Roza, M., Aldeman, C., Swanson, J., & Silberstein, K. (2021, May 18). *30-Minute webinar: Federal funds are flowing! Here's what we're finding so far* [PowerPoint slides]. Edunomics Lab, Georgetown University. Retrieved from https://edunomicslab.org/wp-content/uploads/2021/05/PulseCheck_5.18.2021_V4.pdf
SAS. (2017). Technical documentation of the 2017 TVAAS analyses. Retrieved January 6, 2021, from https://www.tn.gov/content/dam/tn/education/data/tvaas/tvaas_technical_documentation_2017.pdf

Schwartz, H. L., Diliberti, M. K., Berdie, L., Grant, D., Hunter, G. P., & Setodji, C. M. (2021). Urban and rural districts showed a strong divide during the COVID-19 pandemic: Results from the Second American School District Panel Survey. RAND Corporation. Retrieved from https://www.rand.org/pubs/research_reports/RRA956-2.html

Schweig, J. D., Kuhfeld, M., Diliberti, M., McEachin, A., & Mariano, L. T. (2022). Changes in school composition during the COVID-19 pandemic: Implications for school-average interim test score use. RAND Corporation. Retrieved from https://www.rand.org/pubs/research_reports/RRA1037-2.html

Snyder, T. D., Dinkes, R., Sonnenberg, W., & Cornman, S. (2019). *Study of the title I, part A grant program mathematical formulas*. National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubs2019/2019016.pdf

Snyder, B. F., & Parks, V. (2020). Spatial variation in socio-ecological vulnerability to COVID-19 in the contiguous United States. *Health Place*, 66. doi:10.1016/j.healthplace.2020.102471

Stelitano, L., Doan, S., Woo, A., Diliberti, M. K., Kaufman, J. H., & Henry, D. (2020). *The digital divide and COVID-19: Teachers' perceptions of inequities in students' internet access and participation in remote learning*. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA134-3.html

Suspend Statewide Assessments for Select Grades, Colo. House B. 1161 (2021).

Thum, Y. M., & Kuhfeld, M. (2020). *NWEA 2020 MAP Growth achievement status and growth norms for students and schools*. NWEA Research Report. Portland, OR: NWEA.

To Suspend the Public School Rating System for the 2020–2021 School Year Under § 6-15-2105, Ark. House B. 1151 (2021).

U.S. Department of Education. (2021c, February 22). U.S. Department of Education releases guidance to states on assessing student learning during the pandemic. [*Press release*]. Retrieved from https://www.ed.gov/news/press-releases/us-department-education-releases-guidance-states-assessing-student-learning-during-pandemic

U.S. Department of Education. (n.d.) *Template assessment waiver*. Retrieved from https://oese.ed.gov/files/2020/03/template-assessmentwaiver-19-20.pdf

U.S. Department of Education, National Center for Education Statistics. (2021a). 2020-21 common Core of Data (CCD) preliminary files (version 0a) [Data file]. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2021122rev

U.S. Department of Education, National Center for Education Statistics. (2021b) 2019-20 Common Core of Data (CCD) universe files [Data file]. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2021150

Zamarro, G., & Camp, A. (2021, April 20). *Returning to in-person learning might be more complex than simply re-opening schools for minority families*. University of Southern California. Retrieved from https://healthpolicy.usc.edu/evidence-base/returning-to-in-person-learning-might-be-more-complex-than-simply-re-opening-schools-for-minority-families/