



Setting Outcome Goals and Standards From a Formal to Functional Definition of Adequacy

Bruce Baker (Rutgers University) | Drew Atchison (AIR) | Caitlin Kearns (AIR) | Jesse Levin (AIR)

JUNE 2020

Setting Standards Based on Outcomes

According to the New Hampshire Revised Statutes Annotated (RSA) Chapter 193-E, an adequate education is one that meets school approval standards across a variety of subjects, including English language arts (ELA) and reading, mathematics, science, social studies, arts education, and others. The New Hampshire elementary school standards for ELA and reading (Ed 306.37), for example, require “systematic and continuous instruction that develops students’ knowledge of language arts, including listening, speaking, reading, writing, and viewing” (Ed 306.37.a.1) and “instruction which emphasizes how to clarify, order, interpret, and communicate experiences through the skillful use of language” (Ed 306.37.a.2) along with nine additional requirements. Although this language is potentially useful for education leaders and practitioners when developing elementary and secondary educational programs, it provides little guidance as to how one might systematically evaluate whether the standards are being adequately met in all schools. The need to translate opaque and vague language found in state constitutions, statutes, and regulations regarding the provision of education and educational goals into more precise and actionable language for accountability purposes is why states also develop subject- and grade-specific content standards, curricula, and assessments and collect data on educational attainment such as graduation rates and college matriculation. It is through assessment results and other measurable outcomes that states can reasonably determine whether schools and districts are, in fact, meeting the educational goals of the state.

Determining the costs of providing an adequate education requires setting some standard for what constitutes adequate outcome levels. For purposes of cost estimation, that standard must be sufficiently precise and based on existing measures for which there are available data. During the past decade, federal initiatives and numerous states have converged on a framing of standards for elementary and secondary schools focused on “college and career readiness” (U.S. Department of Education, n.d.). This policy emphasis arises from studies in years prior that stressed the lack of preparedness of U.S. high school graduates for college-level work, leading to excessive numbers of students placed in remedial college coursework. The need for remediation has also been accompanied by a marked drop in our nation’s global competitiveness with respect to college completion rates. Thus, to a large extent, college and career readiness standards have been developed with a focus on mitigating this particular problem,

ensuring that high school graduates can succeed in college-level coursework and achieve timely completion of postsecondary degrees (National Research Council, 2008).

Standard setting for college and career readiness has two key components: curricular frameworks and assessments. Developing curricular frameworks involves identifying what high school graduates should know and be able to do to succeed in college-level coursework, as laid out in the Common Core State Standards, and then working backward to develop grade-level standards that define the knowledge and skills students should gain throughout their elementary and secondary school experience.¹ These standards should then guide the development of assessments of student proficiency (Polikoff, Porter, & Smithson, 2011). Both curricular frameworks and assessments are designed to be predictive of college and career readiness; that is, students who have successfully completed these curricular requirements, as evidenced on the assessments, should be more likely to succeed in college-level coursework, persist in postsecondary education, and complete postsecondary degrees. New Hampshire, like many other states, has adopted “College and Career Ready Standards” (Common Core State Standards Initiative, 2020).

On the assessment side of this puzzle—measuring student preparedness—two separate consortium assessments emerged to test whether students were proficient in areas identified in the Common Core State Standards: the Partnership for Assessment of Readiness for College and Careers (PARCC) and the Smarter Balanced Assessment Consortium (SBAC). New Hampshire elected to use the SBAC.

Again, much of the interest in college and career readiness and the Common Core State Standards arose from concerns that students were graduating from high school and achieving “proficiency” on state assessments but were unable to succeed in entry-level coursework in public colleges and universities such as the City University of New York system (Cooper, Cilo, & Baker, 1999). Research by the College Board showed that students scoring 1150 or higher on the SAT were more likely to succeed in first-year college coursework (Kobrin, 2007; Wiley, Wyatt, & Camara, 2011). Building on this work, several researchers have proposed a statistical definition of college readiness using standardized test scores and other measures to predict college success (Fina, Dunbar, & Welch, 2018; Wiley et al., 2011). For example, Wiley et al. (2011) defined college readiness as having at least a 65% probability of achieving a college grade point average of B- or higher, based on measures of high school students’ SAT scores, high school grade point average, and an index of the academic rigor of courses taken.

Statistical definitions of college readiness have had a key influence on the design of modern state assessments as well as the assignment of cut scores intended to denote proficiency for these assessments. Therefore, the equating of student achievement on test scores to college and career readiness can also guide how outcome measures are set for the purpose of modeling the costs of achieving desired levels of educational outcomes. Prior to the Common Core State Standards and the emergence of college and career readiness standards, most states never assessed the validity of chosen cut scores when developing their state assessments; that is, they never asked the following:

Does making the cut lead to greater likelihood of some bigger goal, like succeeding in college level coursework? (Camara, 2013)

¹ See the Common Core State Standards Initiative: <http://www.corestandards.org/>.

Prior to switching to PARCC, New York State engaged Daniel Koretz, PhD, of the Harvard Graduate School of Education to conduct an analysis of their state assessments, including the identification of cut scores that would better reflect college readiness. Dr. Koretz took an approach similar to that of researchers at the College Board, assessing what scores on state assessments were associated with achieving 1150 or higher on the SAT and eventually succeeding in college-level coursework (New York State Education Department, 2010). The consortium assessments SBAC and PARCC established their cut scores on similar bases, with the intent that students scoring proficient or higher would have a significantly higher likelihood of successfully completing college-level coursework (Smarter Balanced Assessment Consortium, 2017).

State assessments can also be measured against the National Assessment of Educational Progress (NAEP), which establishes cut scores for “basic” and “proficient” achievement levels (National Center for Education Statistics [NCES], 2019). The NAEP standards tend to be quite high relative to most state cut scores and even relative to the new SBAC and PARCC assessments (Rahman, Bandeira de Mello, Fox, & Ji, 2019). SBAC standards also differ from PARCC standards. For example, statistical analysis by the American Institutes for Research (Phillips, 2016) found:

1. Smarter Balanced college-ready standards (Level 3) are comparable in difficulty to the NAEP basic levels.
2. Smarter Balanced college-ready standards (Level 3) are significantly below PARCC college-ready standards (Level 4) by about one quarter of a standard deviation.

However, the NAEP standards were not set statistically to represent “college readiness.” As noted by Phillips (2016), NAEP’s proficiency standards are high, at least with respect to SBAC. Alternatively, it may be that SBAC standards are too low, given that SBAC college readiness standards are also lower than PARCC standards.

Setting New Hampshire Standards

Adequacy, while seemingly an absolute construct, also depends on the performance of peers within a given context (Koski & Reich, 2006). In other words, achieving some absolute level of proficiency is important, but context also plays an important role in determining adequacy. Although being academically prepared to pass college-level coursework or take on entry level employment does entail meeting some absolute standards, succeeding in college admissions, or succeeding in the labor market as either a high school or college graduate, also depends on one’s relative position among his or her peers. The relative position of New Hampshire’s students matters, whether it be regionally, nationally, or globally. New Hampshire is situated in a region where individual states (as well as the region collectively) far surpass average national standards and achieve relatively high scores on NAEP. Consequently, benchmarking state outcomes against national averages obscures the region-specific competitive pressures faced by New Hampshire schools and students.

Further, both adequacy and the costs of achieving adequacy depend on spending and outcomes in adjacent states. For example, changes to spending and teacher compensation in Massachusetts affect the cost of labor in southern New Hampshire. Similarly, improved outcomes for children in neighboring

states puts pressure on graduates from New Hampshire schools in terms of access to colleges and universities or labor market opportunities.

New Hampshire's regional position aids in setting standards for cost model estimation precisely because the state's current outcomes are relatively high compared to national averages, as are those of adjacent states. It is much harder and more speculative to project the costs of raising student achievement in Mississippi, for example, where student performance is far below national outcomes and much farther below the level of achievement in high-performing states. In other words, it is much easier to estimate cost models where a sizable portion of schools or districts already perform at levels deemed adequate. Here, we provide several figures that contextualize New Hampshire students' performance relative to NAEP proficiency standards, NAEP equivalent scores for SBAC (Rahman et al., 2019), and neighboring states.

Figure 1 shows that New Hampshire and neighboring states' outcomes on NAEP assessments climbed to their highest levels around 2011 to 2013, falling off somewhat in the years after. For Grade 4 math (top left), average scores in Massachusetts and New Hampshire exceeded the NAEP score associated with SBAC college readiness from 2005 forward, and both exceeded the NAEP proficiency level from 2009 to 2015. Maine and Vermont also exceeded the SBAC proficiency level for several years.

Grade 8 math proficiency levels for both SBAC and NAEP (top right) appear slightly more challenging to achieve than Grade 4 math proficiency levels. Massachusetts has exceeded the SBAC proficiency levels for all years since 2005, while New Hampshire exceeded SBAC proficiency levels between 2007 and 2017. Massachusetts approached the NAEP standard over this time and even exceeded it in one year (2013). The difference between Massachusetts and New Hampshire performance is wider for Grade 8 math than for Grade 4 math.

For Grade 4 reading (bottom left), three states achieved above SBAC college readiness but below NAEP proficiency levels through 2017, with only Maine falling below the SBAC proficiency level. For Grade 8 reading (bottom right), all four states exceeded the SBAC proficiency level through 2017, but none achieved the NAEP proficiency level at any point during the time period. The NAEP proficiency standards for Grade 4 and Grade 8 reading exceed even the average performance of children in Massachusetts, suggesting these standards are quite high.

Given these findings, an adequate outcome goal for New Hampshire children in both the absolute and relative sense might be selected somewhere between the SBAC college readiness and NAEP proficiency standards. The average among the states included here would fall in that range and is similar to New Hampshire's current average performance, if not slightly lower. A marginally higher standard to which the state might wish to strive is the average performance of Massachusetts, which (a) performs marginally higher than New Hampshire on national assessments and (b) exceeded the NAEP proficiency level in some years in math.²

² As mentioned in the research brief on using cost modeling, setting higher outcome standards will also result in a higher adequacy cost (Baker, Atchison, Levin, & Kearns, 2020).

Figure 1. NAEP Performance of New Hampshire and Other New England States Relative to NAEP and SBAC Proficiency Levels

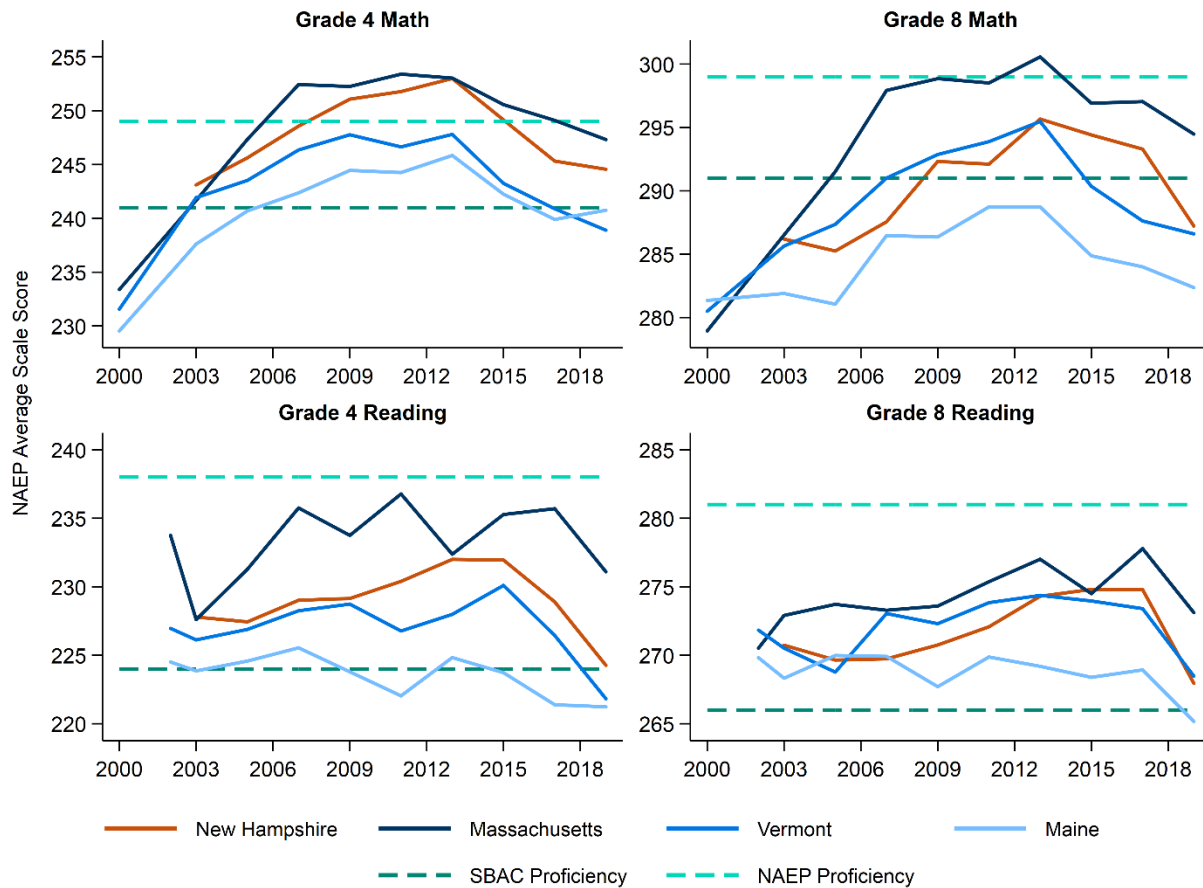
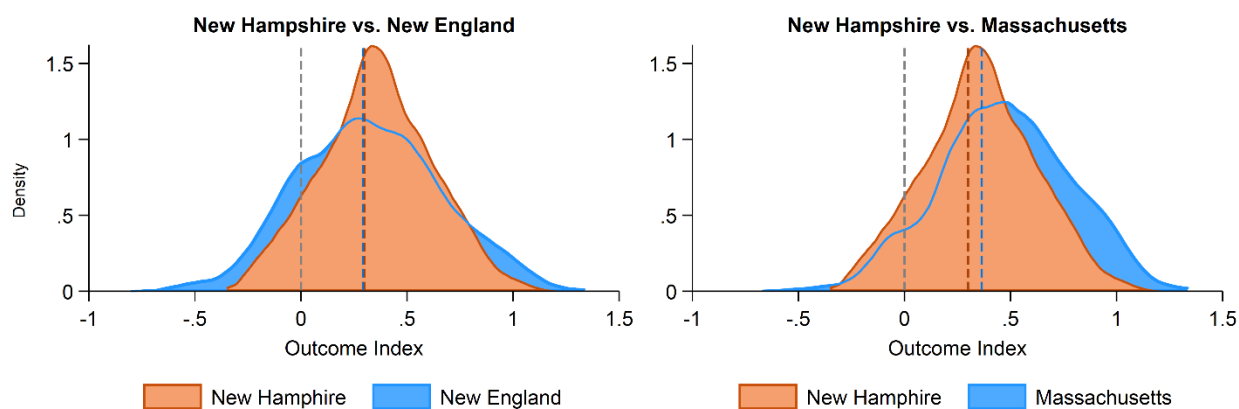


Figure 2 shows the distribution of outcomes for New Hampshire children (a) relative to the average for Maine, Massachusetts, Vermont, and Rhode Island and (b) relative to Massachusetts, specifically. These distributions are based on data from the Stanford Education Data Archive on district-level test scores and are expressed in standard deviations over or under national means, where the national mean is represented as zero. These scores are created for all schools and districts using state assessment data by rescaling and aligning state assessment scores with NAEP, which equates scores across states. The vertical gray dashed line at zero on the x axis represents the national average. Notably, most New Hampshire school districts have scores that exceed the national average by over one-quarter of a standard deviation (just as the above graphs show that New Hampshire children generally score well on NAEP).

The left-hand panel of Figure 2 shows that the average score of all New England districts is approximately equal to the average for New Hampshire (shown by the almost overlapping blue and orange dashed lines); in the right-hand panel, however, the average score for Massachusetts (blue dashed line) is slightly above New Hampshire's (orange dashed line). Further, the Massachusetts distribution in the right-hand

panel is shifted to the right, indicating that outcomes for most school districts in Massachusetts exceed the outcome for the typical New Hampshire school district, echoing the average NAEP scores in the previous graphs.

Figure 2. Distribution of Performance in New Hampshire School Districts Relative to Other New England States and Massachusetts



Notes: The set of states in New England includes Massachusetts, Maine, Rhode Island, and Vermont. The vertical dashed gray line at 0 represents national average outcomes. The vertical dashed orange line represents the enrollment weighted average outcome index for New Hampshire. The vertical dashed blue line represents the enrollment weighted average outcome index for New England and Massachusetts, respectively. In the New Hampshire versus New England figure (left panel) the New Hampshire and New England average outcome indexes are almost identical.

Setting Targets in Kansas

Although some states have relied on input-based models to set basic education funding levels, relatively few states have used rigorous cost modeling to guide their state school finance formulas. In the early 2000s, researchers participating in the Texas School Finance Project provided cost model estimates to the Texas Legislature for their consideration but with negligible influence on the state school finance system (Gronberg, Jansen, Taylor, & Booker, 2005). The first round of Getting Down to Facts in California tested but disregarded cost model estimates, and the eventual reforms adopted years later were not anchored to any input- or outcome-oriented cost estimation (Loeb, Bryk, & Hanushek, 2008).

The Kansas Legislature, however, has sought two separate, independent cost modeling analyses to guide the design of remedies to ongoing school finance litigation. These studies came about in part because of the state’s unique constitutional structure. In Kansas, the State Board of Education (SBOE) and Legislature have independent constitutional authority over public education. SBOE establishes standards under its constitutional responsibility for “general supervision of public schools,” and the legislature is obligated to

“make suitable provision for finance of the educational interests of the state.”³ More background on Kansas’s educational charge is provided in the brief examining finance reform in that state.

How these roles relate to the setting of standards for purposes of determining costs, and defining the legislature’s obligation, came to light in a split between parties defending the state in a constitutional challenge over the adequacy of school funding (*Montoy v. State* [2003]). In response to a Kansas Supreme Court order, the legislature proposed a plan to identify and cost out only the minimal collection of inputs required to provide the core curriculum outlined in SBOE’s standards. Counsel for SBOE, also a defendant, pointed out in their oral arguments to the high court that outcome standards are also under the board’s authority and that any constitutional remedy must consider the costs of achieving outcome standards adopted by the board. The court agreed that the legislature’s funding formula must account for both core curriculum costs *and* the costs of reaching the board’s outcome standards.

The legislature complied by directing the legislative research arm—Legislative Division of Post Audit (LDPA)—to conduct and oversee studies of both the input-based costs and outcome-based costs of achieving standards adopted by SBOE. This framing of constitutional authority and demands that funding be tied to outcome goals has held up through subsequent school funding challenges in Kansas (*Gannon v. State* [2019]). In this brief, we focus on the outcome-based modeling used in the studies that followed these court cases.

Legislative Division of Post Audit Studies With Duncombe and Yinger (2006)

While LDPA in-house researchers conducted the summing of input-based costs, the research component estimating outcome-based costs was contracted out to William Duncombe, PhD, and John Yinger, PhD, of the Maxwell School of Citizenship and Public Affairs at Syracuse University. In collaboration with LDPA staff and under advisement of SBOE, Duncombe and Yinger (a) fit a cost model relating Kansas general fund expenditures and outcomes and (b) projected the per-pupil costs of meeting outcome standards defined as follows:

We use the performance outcomes set by the Kansas State Board of Education for the three math exams, the three reading exams, and the graduation rate. To construct a performance standard comparable to the outcome index used in the cost model, we took a simple average of the standards for these seven performance measures.” (Duncombe & Yinger, 2005, p. C-33)

Researchers in LDPA combined findings of Duncombe and Yinger’s model with their own input-based analysis to suggest cost targets for each district, which the legislature would then be responsible for funding (Kansas Legislative Division of Post Audit, 2006). This approach would meet the court’s demand that the funding targets take into account outcome standards established by SBOE.

One shortcoming with this work is that these standards had never been externally validated. Test developers at the University of Kansas had not performed the kinds of predictive validity checks described in the first section of this brief, which evaluate whether cut scores are reasonably predictive of postsecondary success. In fact, in the first few iterations of the NCES *Mapping Standards* reports, which

³ See Kansas Constitution, Article 6: <https://kslib.info/832/Article-6-Education>.

compare the rigor of proficiency targets across states, Kansas cut scores were found to be especially low when aligned with NAEP standards for proficiency (National Center for Education Statistics, 2009).

WestEd/Taylor Study (2018)

Twelve years later, Kansas was finally digging its way out from recession era tax cuts and lagging revenues while facing a judicial order (*Gannon v. State [2018]*) to adequately fund schools. Legislators were once again looking to reestimate the costs associated with meeting their constitutional obligation. Wishing to maintain a similar air of independence, rigor, and credibility, the legislature again contracted with an outside expert, this time Lori Taylor, PhD, of Texas A&M University, to estimate a cost model to determine the cost of fulfilling their obligation (Taylor, Willis, Berg-Jacobson, Jaquet, & Caparas, 2018). Dr. Taylor and her team were given sufficient data and independent access to work with SBOE to define standards for use in cost projections.

Ultimately, Taylor et al. (2018) chose to define adequacy using graduation rates and student assessment scores on Kansas state standardized tests. The authors motivate their use of these measures by referencing the standards generated by the influential Kentucky State Supreme Court ruling in the *Rose v. Council of Better Education* case (1989). Specifically, Dr. Taylor’s team linked Kansas standardized tests and graduation rates to the seven Rose standards, which define the capacities that all students are expected to achieve:

- *Sufficient oral and written communication skills to enable them to function in a complex and rapidly changing civilization;*
- *Sufficient knowledge of economic, social, and political systems to enable them to make informed choices;*
- *Sufficient understanding of governmental processes to enable them to understand the issues that affect their community, state, and nation;*
- *Sufficient self-knowledge and knowledge of their own mental and physical wellness;*
- *Sufficient grounding in the arts to enable them to appreciate their cultural and historical heritage;*
- *Sufficient training or preparation for advanced training in academic or vocational fields, to enable them to choose and pursue life work intelligently;*
- *Sufficient academic or vocational skills to enable them to compete favorably with their counterparts in surrounding states, in academics or in the job market. (Rose v. Council for Better Education, 790 S.W.2d 186, 60 Ed. Law Rep. 1289, 1989)*

Dr. Taylor’s analysis adopted two scenarios for the cost estimates. These scenarios both set a graduation rate of 95% but identified different performance thresholds for student assessments (see Table 1). Scenario A is described as “approaching on track for college readiness” and makes use of the percentage of students meeting Level 2 on the Kansas Assessment Program (KAP) ELA/math tests. The more

stringent Scenario B is described as “on track for college readiness” and uses the percentage of students scoring at Level 3 or above on the same assessments.⁴

Table 1. Percentage of Students Meeting Performance Thresholds Under Two Different Scenarios

	ELA (Level 2+)	Math (Level 2+)	ELA (Level 3+)	Math (Level 3+)	Graduation Rate
Scenario A	90	90			95
Scenario B			60	60	95

Source: Table 15 on page 64 from Taylor et al. (2018).

Taylor and colleagues (2018) explain in their report that Scenario A approximates the expected performance levels under former assessments during the No Child Left Behind era. These are the standards that were in place at the time of the Duncombe and Yinger (2005) and Kansas LDPA (2006) studies. Scenario B is based on Kansas’s updated college and career readiness standards and uses a more stringent performance threshold.

To put the different standards used in the Kansas study into perspective, Table 2 shows for 4th and 8th grade reading and math: (a) NAEP cut scores denoting proficiency, (b) SBAC proficiency cut scores equated to NAEP, (c) 2009 Kansas proficiency cut scores equated to NAEP, and (d) 2017 Kansas proficiency cut scores equated to NAEP. The figures show that Kansas proficiency standards in 2009 were quite low, aligned with “basic” levels or below on NAEP and surpassed by most children in the state. Kansas’s updated “college and career ready” assessment standards are more stringent, especially at the 8th grade level, and current proficiency rates are therefore much lower. Therefore, raising the rates of children surpassing these higher standards (cut scores) is a more difficult task requiring a larger investment of resources. In turn, the projected cost of providing an adequate education under Scenario B was higher than for Scenario A despite setting a lower percentage of students meeting the more stringent 2017 standards (at 60% instead of 90%).

Table 2. Relationship Between NAEP, SBAC and Kansas Assessments, Then and Now

Test	NAEP 2017	SBAC 2017	Kansas 2009	Kansas 2017
Reading Grade 4	238	224	186	228
Reading Grade 8	281	266	236	290
Math Grade 4	249	241	217	251
Math Grade 8	299	291	265	310

Notes: For mapping of Kansas 2009 standards to NAEP, see National Center for Education Statistics (2011). For mapping of Kansas 2017 standards to NAEP, see Rahman et al. (2019).

⁴ The KAP performance levels range from 1 to 4 as follows: At Level 1, a student is below grade level; at Level 2, a student is at grade level but not on track for college or career readiness; at Level 3, a student is at grade level and on track for college or career readiness; and, at Level 4, a student exceeds grade-level expectations and is on track for college or the workplace.

Summary

The lesson to be learned from Kansas is that outcome standards matter for estimating costs. The passage of more rigorous college and career ready standards raised educational expectations, and these more rigorous educational goals were found to require increased investment to achieve adequacy.

Perhaps more importantly, tested outcomes and graduation rates are intermediate checkpoints in the system, not the end goals themselves. They are valuable only to the extent that they have some validity in predicting longer term outcomes more valuable than simply doing well on a test. These outcomes might include successfully completing college-level coursework and persisting to degree or certificate completion.

Therefore, to the extent that assessments are used in modeling costs, we suggest two guiding principles:

- First, that assigned outcome levels have some relation to external goals, based on evaluation of the predictive validity of the assessments used in relation to those goals.
- Second, that assigned outcome levels take into account the relative position of a state's children in a global, national, and regional context.

The first of these speaks directly to constitutional obligations, where judicial interpretations of state constitutions often speak to the right of children to become sufficiently informed and engaged citizens. The second of these principles acknowledges the necessarily relative nature of educational outcomes, and that students from New Hampshire must find themselves on a level playing field with neighboring states, the nation, and ultimately the global economy. In this sense, New Hampshire is fortunate to be among peer states that sit at the top in terms of national educational achievement, and that when considered globally compare favorably against top nations (Baker & Weber, 2016).

References

- Baker, B., Atchison, D., Levin, J., & Kearns, C. (2020). *Using cost modeling to inform education funding formulas*. Washington, DC: American Institutes for Research.
- Baker, B. D., & Weber, M. (2016). *Deconstructing the myth of American public schooling inefficiency*. Washington, DC: Albert Shanker Institute.
- Camara, W. (2013). Defining and measuring college and career readiness: A validation framework. *Educational Measurement: Issues and Practice*, 32(4), 16–27.
- Common Core State Standards Initiative. (2020). *Standards in your state*. Retrieved from <http://www.corestandards.org/standards-in-your-state/>
- Cooper, B. S., Cilo, M. R., & Baker, B. D. (1999). Making the transition from school to college: The case of New York City Public Education. In M. Alampi & P. M. Comeau (Eds.), *American education annual: Trends and issues in the educational community (1998-1999)*. New York, NY: Gale Research
- Duncombe, W., & Yinger, J. (2005). *Estimating the costs of meeting student performance outcomes adopted by the Kansas State Board of Education*. Syracuse, NY: Syracuse University. Retrieved from https://www.maxwell.syr.edu/uploadedFiles/cpr/research/cpr_research_education_finance_policy/Kansas_Report.pdf
- Fina, A. D., Dunbar, S. B., & Welch, C. J. (2018). Establishing empirical links between high school assessments and college outcomes: An essential requirement for college readiness interpretations. *Educational Assessment*, 23(3), 157–172.
- Gronberg, T., Jansen, D., Taylor, L., & Booker, K. (2005). *School outcomes and school costs: The cost function approach*. College Station, TX: Texas A&M University. Retrieved from <https://bush.tamu.edu/research/faculty/TXSchoolFinance/papers/SchoolOutcomesAndSchoolCosts.pdf>
- Guthrie, J. W., & Rothstein, R. (1999). Enabling “adequacy” to achieve reality: Translating adequacy into state school finance distribution arrangements. In *Equity and adequacy in education finance: Issues and perspectives* (pp. 209–259). Washington, DC: The National Academies Press.
- Kansas Legislative Division of Post Audit. (2006). *Cost study analysis—Elementary and secondary education in Kansas: Estimating the costs of K–12 education using two approaches*. Topeka, KS: Author. Retrieved from http://www.kslegresearch.org/KLRD-web/Publications/Education/Education_Cost_Study/Cost_Study_Report.pdf
- Kobrin, J. (2007). *Determining SAT benchmarks for college readiness* (Research Notes RN-30). New York, NY: College Board, Office of Research and Analysis. Retrieved from <https://files.eric.ed.gov/fulltext/ED562605.pdf>
- Koski, W. S., & Reich, R. (2006). When adequate isn’t: The retreat from equity in educational law and policy and why it matters. *Emory Law Journal*, 56(3), 545–617.
- Loeb, S., Bryk, A., & Hanushek, E. (2008). Getting down to facts: School finance and governance in California. *Education Finance and Policy*, 3(1), 1–19.

- National Center for Education Statistics. (2009). *Mapping state proficiency standards onto the NAEP scales: Variation and change in state standards for reading and mathematics, 2005–2009: Kansas*. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/nationsreportcard/pdf/studies/mapping_factsheets/2009_Mapping_Factsheets_KS.pdf
- National Center for Education Statistics. (2011). *Mapping state proficiency standards onto the NAEP scales: Variation and change in state standards for reading and mathematics, 2008–2009*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved from <https://nces.ed.gov/nationsreportcard/pdf/studies/2011458.pdf>
- National Center for Education Statistics. (2019). *The NAEP mathematics achievement levels by grade*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved from <https://nces.ed.gov/nationsreportcard/mathematics/achieve.aspx>
- National Research Council. (2008). *Common standards for K–12 education?: Considering the evidence: Summary of a workshop series*. Washington, DC: National Academies Press.
- New York State Education Department. (2010). *Statement on research related to proficiency on state assessments*. Retrieved from http://usny.nysed.gov/scoring_changes/
- Phillips, G. (2016). *National benchmarks for state achievement standards*. Washington, DC: American Institutes for Research. Retrieved from https://www.air.org/sites/default/files/downloads/report/National-Benchmarks-State-Achievement-Standards-February-2016_rev.pdf
- Polikoff, M. S., Porter, A. C., & Smithson, J. (2011). How well aligned are state assessments of student achievement with state content standards? *American Educational Research Journal*, 48(4), 965–995.
- Rahman, T., Bandeira de Mello, V., Fox, M. A., & Ji, C. S. (2019). *Mapping state proficiency standards onto the NAEP scales: Results from the 2017 NAEP reading and mathematics assessments* (NCES 2019-040). Washington, DC: National Center for Education Statistics.
- Smarter Balanced Assessment Consortium. (2017). *Smarter Balanced cut score validation: Final report*. Santa Clara, CA: Author. Retrieved from <https://portal.smarterbalanced.org/library/en/smarter-balanced-cut-score-validation-final-report.pdf>
- Taylor, L., Willis, J., Berg-Jacobson, A., Jaquet, K., & Caparas, R. (2018). *Estimating the costs associated with reaching student achievement expectations for Kansas public education students: A cost function approach*. San Francisco, CA: WestEd. Retrieved from https://kasb.org/wp-content/uploads/2018/03/Kansas_Adequacy-Study_Cost-Function_20180315FINAL_02.pdf
- U.S. Department of Education. (n.d.). *College- and career-ready standards*. Washington, DC: Author. Retrieved from <https://www.ed.gov/k-12reforms/standards>
- Wiley, A., Wyatt, J., & Camara, W. J. (2011). *The development of a multidimensional college readiness index* (Research Report 2010-3). New York, NY: College Board.



AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

About the American Institutes for Research

Established in 1946, with headquarters in Washington, D.C., the American Institutes for Research (AIR) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance, both domestically and internationally, in the areas of education, health, and the workforce. **For more information, visit www.air.org.**

Notice of Trademark: "American Institutes for Research" and "AIR" are registered trademarks. All other brand, product, or company names are trademarks or registered trademarks of their respective owners.

Copyright © 2020 American Institutes for Research®. All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, website display, or other electronic or mechanical methods, without the prior written permission of the American Institutes for Research. For permission requests, please use the Contact Us form on www.air.org.