

case will be dismissed with prejudice as to all defendants other than the Anderson County Schools Board of Education (the “ACSBE”).¹

BACKGROUND

I. The First to the Top Act and the TVAAS

A. The Act

In January 2010, Tennessee enacted the “First to the Top Act” (the “Act”) which, among other things, changed the way that Tennessee public school teachers are evaluated.² This case involves a challenge under the United States Constitution to the implementation of that Act at the state and local levels.

The Act has been amended several times, but one of its basic features has remained constant: the Act requires that teachers be evaluated in part based on measures of “student achievement data,” including (1) “student achievement data” generated by the Tennessee Value-Added Assessment System (“TVAAS”) or some “other comparable measure of student growth” and (2) “other measures of student achievement.” Tenn. Code Ann. § 49-1-302(B)(ii)-(iii) (as of April 15, 2015).³

¹ In response to the court’s June 3, 2015 Order (Docket No. 55), the ACSBE informed the court that it was not joining either of the pending motions (Docket No. 56).

² Including subsequent amendments, the Act has been codified as Tenn. Code Ann. §§ 49-1-302 *et seq.*, 49-1-602 *et seq.*, and 49-5-506 *et seq.*

³ The Act has been amended multiple times since it was passed into law in January 2010, including recent amendments effective April 16, 2015. The plaintiffs largely gloss over these distinctions, and neither party has acknowledged the sweeping legislative amendments that went into effect on April 16, 2015, which have materially changed many of the provisions implicated in this lawsuit. To keep things clear, unless otherwise noted, the court will refer to the Act as it was in place when the plaintiffs filed this lawsuit in February 2015. In some instances, which the court will note herein, the court will refer to the Act as it was enacted in January 2010 or as it

The Act contemplates that these performance measures must include some form of value added assessment (“VAA”), which the Act defines as “[a] statistical system for educational outcome assessment that uses measures of student learning to enable the estimation of teacher, school, or district statistical distributions[.]” *Id.* § 603(a)(1). The Act requires that the VAA “will use available and appropriate data as input” to estimate the “impact” or “effect” that a teacher, school, or school district has “on the educational progress of students.” *Id.* § 603(a)(2). The Act also states that any VAA utilized to measure the performance of Tennessee teachers, schools, and school districts must “cover[] the total range of topics covered in the approved curriculum” and “should have [a] strong relationship to the core curriculum for the applicable standard grade level and subject.” *Id.* § 603(c).⁴

was amended on April 16, 2015 by the Tennessee Teacher Evaluation Enhancement Act and other enacted bills. *See* TN LEGIS 182 (H.B. 832) (approved Apr. 16, 2015); TN LEGIS 158 (H.B. 108) (approved Apr. 16, 2015); and TN LEGIS 304 (S.B. 341) (approved Apr. 24, 2015). In general, all references to the Act will cite only to a section or subsection within Chapter 1 of Title 49 of the Tennessee Code Annotated (“§ 302(d),” for example) without repetitively citing that chapter and title.

⁴ In scholarly literature incorporated by reference into the Complaint, a “VAA” is also referred to as a “value-added metric,” or “VAM.” *See* American Statistical Association Statement on Using Value-Added Models for Educational Assessment (Apr. 8, 2014) (referenced in ¶ 28 of the Complaint), *available at* http://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf (last accessed June 12, 2015). As the American Statistical Association explains in this statement:

Many states and school districts have adopted Value-Added Models (VAMs) as part of educational accountability systems. The goal of these models, which are also referred to as Value-Added Assessment (VAA) Models, is to estimate effects of individual teachers or schools on student achievement while accounting for differences in student background. VAMs are increasingly promoted or mandated as a component in high-stakes decisions such as determining compensation, evaluating and ranking teachers, hiring or dismissing teachers, awarding tenure, and closing schools.

The Act created a fifteen-person “teacher evaluation advisory committee” (the “TEAC”). *Id.* § 302(d)(1). By statute, the TEAC is composed of a diverse group of people, including the Tennessee Commissioner of Education, the executive director of the State Board of Education (the “State Board”), the chairpersons of the Tennessee Senate and House of Representatives education committees, one K-12 public school teacher appointed by the speaker of the Senate, one K-12 teacher appointed by the speaker of the House, and nine members appointed by the Governor, including three public school teachers, two public school principals, one director of a school district, and three members representing “other stakeholders” interests. *Id.* The membership of this committee must “appropriately reflect the racial and geographic diversity” of Tennessee, and at least one member of the committee must be a parent of a currently enrolled public school student. *Id.*

. . . VAMs attempt to predict the “value” a teacher would add to student achievement growth, as measured by standardized test scores, if each teacher taught comparable students under the same conditions. VAM results are often regarded as more objective or authoritative than other types of information because they are based on student outcomes, use quantitative complex models, and rely on standard test scores and common procedures for all teachers or schools. . . .

Value-added models typically use a form of regression model predicting student scores or growth on standardized tests from background variables (including prior test scores), with terms in the model for the teachers who have taught the student. The model coefficients for the teachers are used to calculate their VAM scores. In related models known as “growth models” a regression model is fit to predict students’ current test scores from previous tests scores. A percentile is calculated for each student from the model, relating his or her growth to the growth of other students with similar previous test scores. The median or average of the percentiles of a teacher’s students is then used to calculate the teacher’s VAM score.

Id. at pp. 1 and 3.

The TEAC was (and remains) responsible for developing and recommending to the State Board “guidelines and criteria for the annual evaluation” of Tennessee’s teachers.

Id. § 302(d)(2)(A). As set forth in § 302(d)(2)(B)(ii), the Tennessee legislature contemplated that the TVAAS would constitute an appropriate form of “student achievement data” for evaluating teachers, provided that it met the criteria set forth in § 603. The Act also provides that the “other measures of student achievement” must include “measures developed by the [TEAC] and adopted by the [State Board],” which “shall be verified by the department of education to ensure that the evaluations correspond with the teaching assignment of each individual teacher” *Id.* § 302(d)(2)(B)(iii).

The Act contains specific provisions relating to the use of testing data that form the basis for the VAA of each teacher’s performance. Section 606 requires that “data from the Tennessee comprehensive assessment program (TCAP) tests, or their future replacements,” be used to evaluate “teacher effects on the educational progress of students within school districts for grades three through eight (3-8).” *Id.* § 606(a). Section 608 requires that “[t]he development of subject matter tests will be initiated to measure performance of high school students in subjects designated by the state board of education and reviewed by the education committee of the senate and the education committee of the house of representatives.” These tests “must cover the complete range of topics covered within the list of state approved textbooks and instructional materials for that subject.” *Id.* “As soon as valid tests have been developed, the testing of students will be initiated to provide for value added assessment[,]” which must be conducted annually. *Id.*

The Act also states that tests in additional subjects should be initiated in the future: “Value added assessment may be initiated in other subjects designed by the state board of

education and reviewed by the education committee of the senate and the education committee of the house of representatives at such times as valid tests can be developed that effectively measure performance in such subjects.” *Id.*

In other words, the Act required that teachers be evaluated in substantial part based on a value-added assessment (the TVAAS, if implemented in time), created a special committee to develop teacher evaluation standards for review and approval by the State Board, and envisioned that additional value-added assessments would continue to be developed and implemented over time, with input from the TEAC (which includes legislative and executive appointees and the Commissioner of Education herself), the State Board, and the relevant legislative committees of both branches of the Tennessee legislature.

B. The TVAAS, Individual Scores, and School-Wide Scores

According to the plaintiffs’ Complaint, the TVAAS is a statistical model owned by a private, for-profit company that has not disclosed its precise methodology. To the plaintiffs’ knowledge, the TVAAS utilizes students’ performance on past standardized tests to project how those students will perform on a subsequent standardized test. The TVAAS methodology compares the students’ actual performance on a recent standardized test to the students’ predicted performance. The differential relative to expectations for that group of students (better, worse, or the same) is attributed in some fashion to the teacher’s impact on that student. The TVAAS generates a teacher evaluation score from 1 to 5, where a “5” purports to indicate that the teacher performed “significantly above expectations” and a “1” purports to indicate that the teacher performed “significantly below expectations.” Under the terms of the Act, the TVAAS score is supposed to reflect “[t]he impact that a teacher . . . has on the progress, or lack of progress, in educational attainment or learning of a student” § 603(a)(2).

The TVAAS relies on testing data relating to just a subset of the subjects offered in Tennessee public schools. For students in grades 3-8, the TVAAS incorporates data from the Tennessee Comprehensive Assessment Program (the “TCAP”), which tests only mathematics, reading/language arts, social studies, and science.⁵ For students in grades 9-12, the TVAAS relies upon certain state-approved high school end-of-course (“EOC”) examinations, which test only Algebra I and II, English I, II, and III; Biology I; Chemistry; and U.S. History.⁶ The court will refer to the particular subjects incorporated into the TVAAS (via the TCAP or the EOC examinations) as “tested subjects.” The court will refer to teachers of these subjects as “teachers of tested subjects.”

As alleged in the Complaint, many (perhaps more than half) of Tennessee’s public school educators teach courses that are not among the specific subjects incorporated into the TVAAS at any grade level. These subjects include, among others, physical education, Spanish, art, world history, physics, computer technology, agriculture, finance, information technology, and hospitality and tourism, to name a few.⁷ The court will refer to these subjects collectively as “non-tested subjects” and will refer to teachers of these subjects as “teachers of non-tested subjects.”

⁵ As stated in the previous section of this opinion, § 606 appears to require the use of the TCAP for teachers of students in grades 3-8.

⁶ For students in grades K-2, districts can also elect for the TVAAS to incorporate data from a K-2 assessment. The plaintiffs’ Complaint and the parties’ briefs do not indicate how K-2 teachers are evaluated when the K-2 assessment is not utilized.

⁷ A publicly available list of Tennessee public school course offerings can be found at <http://tn.gov/education/standards/> (last visited June 12, 2015).

The plaintiffs allege that the TVAAS can generate two types of scores. The first is an individual score, which measures how the students of a teacher of a tested subject have performed on standardized exams that test – at least in part – that teacher’s coursework. As the court understands it, the individual score varies by teacher, depending on the mix of students taught by that teacher.

The second type of TVAAS score is a school-wide composite score. That school-wide score measures how, relative to expectations, *all* of the students at a school performed on the standardized tests that are incorporated into the TVAAS estimate (either the TCAP or the EOC examinations, depending on the grade level). The school-wide TVAAS score is therefore an aggregated, unitary score that does not vary by teacher.

C. The Plaintiffs’ Criticisms of VAAs and the TVAAS

The plaintiffs allege that, as a general matter, VAAs are severely flawed assessment tools for evaluating an individual teacher’s performance. They allege that VAAs fail to control adequately for various factors outside the control of an individual teacher that can impact test scores, including out-of-school factors (such as parental support for learning, poverty, and English learner status) and school-wide factors (such as class size, instructional time, and school resources). The plaintiffs also point out that VAAs are the subject of scholarly criticism. For example: (1) in October 2009 (shortly before the Act was passed), the National Academy of Sciences urged that VAAs should not play a “significant” role in “high-stakes teacher evaluation systems;” (2) in August 2010, shortly after the Act was passed, one study concluded that only 4% to 16% of variation in a teacher’s value-added ranking could be predicted from that teacher’s rating the previous year; and (3) in April 2014 (over four years after the Act was passed), the

American Statistical Association issued a statement that using VAAs to measure teacher performance “can have unintended consequences that reduce quality.”

The plaintiffs allege that the TVAAS, like all VAAs, suffers from these infirmities with respect to teachers of tested subjects who receive an individual score. With respect to the TVAAS itself, the plaintiffs also allege that each raw TVAAS score has a high margin of error that spans multiple number ratings. For example, within a reasonable confidence interval, a raw “3” score might actually reflect a “2” or a “4” level of performance. In light of this variability, the plaintiffs allege that it is problematic to report the score as a single number.

Although the plaintiffs severely criticize using the TVAAS to measure the performance of teachers of tested subjects, they do not assert a constitutional challenge to the use of individual TVAAS scores for that purpose. Instead, they challenge the constitutionality of the State Board’s policy (described herein) that *school-wide* TVAAS scores must be used to evaluate the performance of teachers of *non-tested subjects*. As to those teachers, the plaintiffs allege that the TVAAS school-wide composite score provides no measure of individual teacher performance whatsoever. With respect to those teachers, the TVAAS school-wide score incorporates both (1) the performance of students that the teacher did not actually teach (because each teacher presumably does not teach every student at the school) and (2) the performance of all the school’s students on subjects that the teacher does not teach. For example, if an excellent high school art teacher teaches 200 of a school’s 1000 students in school year X, the TVAAS school-wide composite for school year X would be based in large part on students not instructed by that teacher (*i.e.*, 200 students taught by that teacher and 800 students not taught by that teacher) and entirely on all students’ performance on subjects other than high school art. Moreover, in a jurisdiction utilizing the school-wide TVAAS score as a component of non-tested subject teacher

evaluations, the excellent art teacher would receive the same TVAAS performance rating as an ineffective computer technology teacher at the same school.

The plaintiffs allege that the TVAAS is not designed to measure the performance of teachers on non-tested subjects. They allege that the school-wide TVAAS score assigns a number rating to teachers of non-tested subjects that is unrelated to the work that the teacher is licensed and required to teach. Finally, they allege that the TVAAS school-wide composite score therefore bears no relationship whatsoever to the quality of instruction provided by a teacher of a non-tested subject.

D. The Act's Percentage Requirements

Since its inception, the Act has required that a VAA comprise some measure of a teacher's performance evaluation.

At the time it was passed – and ultimately effective for the 2011-12 and 2012-13 school years – the Act required that the TEAC develop, and the State Board approve, criteria for the evaluation of teachers based 50% on “student achievement data.” *See* § 302(d)(2)(A) and 302(d)(3) (as enacted Jan. 16, 2010) (stating that implementation would be effective no later than the 2011-12 academic year). Of this 50%, the Act initially required that (1) 35% of the evaluation criteria “be student achievement data based on student growth data as represented by the [TVAAS] . . . or some other comparable measure of student growth, if no such TVAAS data is available,” and (2) 15% of the evaluation criteria be “based on other measures of student achievement data selected from a list of such measures developed by the [TEAC] and adopted by the board.” *Id.* § 302(d)(2)(A)(i)-(ii) (as enacted Jan. 16, 2010). As to the latter student achievement component, the Act indicated that the evaluator and evaluated teacher should try to agree on a selection from that “list” of “other” options, but that the evaluator's choice would

prevail, should they disagree. *Id.* § 302(d)(2)(A)(ii) (as enacted Jan. 16, 2010). The Act did not specify what information should comprise the remaining 50% of the evaluation, other than to indicate that it should be based on teacher-specific observational data. *Id.* § 302(d)(2)(B) (as enacted Jan. 16, 2010).

Among other amendments to the Act, effective April 11, 2013, the legislature amended the section relating to the evaluation criteria percentages for student achievement data – *see* 2013 Tenn. Laws Pub. Ch. 105, TN LEGIS 105 (H.B.150) (2013) – and those changes remained in effect through April 15, 2015. In most relevant part, the legislature changed the evaluation percentages for “teachers without access to individual data representative of student growth[.]” *Id.* § 302(d)(2)(B)(vi) (as amended through April 15, 2015).⁸ For those teachers, only 40% of the evaluation (rather than 50%) is comprised of “student achievement data,” including just 25% (rather than 35%) “based on student growth data as represented by the TVAAS or some other comparable measure if no such TVAAS data is available.” *Id.* Furthermore, as to the remaining 15% (based on “other measures of student achievement selected from a list” of measures developed by the TEAC and approved by the Board), the Act was amended to allow the evaluated teacher’s selection to prevail over the evaluator’s choice, in the event of a disagreement.

In other words, effective for the 2013-14 school year, the legislature made two changes that impacted the evaluations of teachers without access to individualized student performance data. First, it acknowledged that data from a VAA developed in compliance with § 302(d) might not reflect individualized data for certain teachers, with respect to whom it accordingly reduced

⁸ This provision was added as § 302(d)(2)(A)(vi) at the time. Under subsequent amendments to the Act, it was relabeled as § 302(d)(2)(B)(vi).

by 10% (from 35% to 25% for one component, and from 50% to 40% overall) the portion of the evaluation that would be based on the VAA (effectively, the TVAAS). Furthermore, as to the remaining 15% of the student achievement component of each teacher’s evaluation, it gave teachers, rather than evaluators, the power to choose alternative metrics from a list of approved student achievement measures.

Both before and after this amendment, the State Board has required that teachers be assigned an overall performance rating of 1 to 5, where a “1” indicates that the teacher performed “significantly below expectations” and a “5” indicates that the teacher performed “significantly above expectations.”⁹ The overall performance rating incorporates (1) a teacher’s performance rating based on observational data, (2) a teacher’s performance rating based on the TVAAS score, and, in some cases, (3) a performance rating based on some “other measure” of student growth other than TVAAS, with separate weight ascribed to each factor (*i.e.*, 50% observational, 25% TVAAS individual or school-wide score, and 15% based on one of the “other measures” other than the TVAAS).

E. Use of Numerical Performance Evaluations

Since its inception, the Act has stated that the teacher evaluations “shall be *a factor* in employment decisions, including, but not necessarily limited to, promotion, retention, termination, compensation, and the attainment of tenure status.” *Id.* § 302(d)(2)(A). However, the Act does not state how much weight should be assigned to this “factor.”

F. Upcoming and Proposed Changes to the Act

⁹ Perhaps to align the TVAAS scores with the Tennessee numerical rating system, the TVAAS also rates teachers on a parallel 1 to 5 scale.

Although not mentioned by the parties here, publicly available records indicate that the Tennessee legislature recently amended the Act and may be considering additional changes to it. On April 16, 2015, Governor Haslam signed into law the Tennessee Teaching Evaluation Enhancement Act (the “TEEA”), which makes significant changes to the Act. *See* TN LEGIS 158 (H.B. 108). Among other changes, it (1) modifies § 302(d)(2)(A) to indicate that school districts are not required to use student achievement data based on state assessments as the sole factor in employment decisions; (2) for teachers *with access* to individualized student growth data, it re-phases the weight assigned to certain components of the VAA for school years 2015-16, 2016-17, and 2017-18; and (3) for teachers *without access* to individualized student growth data, a TVAAS score or “some other comparable measure” may only comprise 10% of a teacher’s evaluation for the 2015-16 school year and only 15% of that teacher’s evaluation thereafter. *See also* TN LEGIS 182 (approved Apr. 16, 2015) (amending, *inter alia*, § 608).

Other amendments have been proposed in recent months, some of which would make even more sweeping changes, if enacted.¹⁰

II. The State Board Policy Implementing the Act

Tennessee law authorizes the existence of a State Board of Education and entrusts it within certain responsibilities, including the responsibilities set forth in the Act. *See* § 301 *et*

¹⁰ *See, e.g.*, 2015 TN. S.B. 802 (Feb. 12, 2015) (proposing to add a provision eliminating the use of a numerical system to evaluate teacher and principal performance in favor categories for “exceeds expectations,” “meets expectations,” and “below expectations.”); 2015 TN H.B. 1130 (Feb. 12, 2015) (proposing sweeping changes to the composition of the TEAC to include, *inter alia*, a member of the Tennessee Education Association and a member of the Tennessee Principals Association, and proposing a moratorium on the use of TVAAS for purposes of principal and teacher evaluations until the newly composed TEAC develops new teacher performance evaluation criteria).

seq. By statute, the State Board is composed of a diverse group of nine members appointed by the Governor and confirmed by the Tennessee Senate and House of Representatives. *See* § 301(a)(1)-(3). The appointees must include one public high school student, one *ex officio* member, at least one appointed member from each of Tennessee’s congressional districts, one member of a minority race, at least three members from each of the majority and minority political parties, and one K-12 public school teacher. *Id.* § 301(a)(1). No appointed board member can be an elected official or employee of the federal, state, or a local government. *Id.* § 301(a)(4). Subject to some qualifications not relevant here, board members generally serve staggered five-year terms. *Id.* § 301(a)(2). The State Board’s meetings, which must occur at least quarterly at the state capital, must be made available (both live and in archived form) for public viewing on the internet. §§ 301(d)(1) and (d)(4).

Among many responsibilities, the State Board must set policies for evaluating individual teachers and for evaluating individual student progress and achievement. § 302(a)(2)(C). The State Board also must adopt policies for evaluating teachers (*id.* § 302(a)(5)(A)(iii)) and policies relating to teacher compensation (*id.* § 302(a)(5)(B)).¹¹ As mentioned in the previous section, the Act requires the Board to consider recommendations from the TEAC and to adopt new teacher evaluation standards premised in part on the TVAAS or another VAA tool, *id.* § 302(d)(2)(A), and the Act requires the State Board to designate VAAs to measure “other subjects . . . at such times as valid tests can be developed that effectively measure performance in such subjects.” *Id.* § 608.

¹¹ The Board has also created rules governing, among other things, teacher evaluations. *See* Rules of the State Board of Education, Ch. 0520-02-01 (last revised May 2012), available at <http://state.tn.us/sos/rules/0520/0520-02/0520-02-01.20120530.pdf> (last accessed June 12, 2015). The rules appear to be distinct from the policies.

At some point following the Act’s passage (and effective for the 2011-12 school year), the State Board received recommendations from the TEAC and adopted a new “Teacher and Principal Evaluation Policy.” See Tennessee State Board of Education Policy 5.201 (the “State Board Policy”).¹² In this case, the plaintiffs’ Complaint attaches one intermediate post-Act iteration of that policy. However, publicly available versions of the policy reflect earlier versions that post-date the Act and additional versions through January 30, 2015 – just one week before the plaintiffs filed this lawsuit.¹³ The earliest publicly available post-Act version of the policy that the court has located was revised November 4, 2011, and stated as follows:

¹² Although not referenced by the parties here, publicly available agendas released by the State Board reflect that, in multiple board meetings, the Board’s agenda indeed included the consideration of recommendations from the TEAC with respect to the initial teacher evaluation policy and with respect to subsequent amendments to that policy. See, e.g., Agenda for January 28, 2011 State Board Meeting, Item IV.F (concerning approval of implementing rules in light of the Act and TEAC recommendations), *available at* <http://www.tn.gov/sbe/2011Januarypdfs/IV%20F%20Evaluation%20&%20Grievance%20Rules%20Cover%20Sheet%20&%20Attachment.pdf> (last accessed June 12, 2015); Agenda for April 15, 2011 State Board Meeting, Item IV.M (concerning approval of educator evaluation policy), *available at* <http://www.tn.gov/sbe/2011Aprilpdfs/IV%20C%20Teacher%20and%20Principal%20Evaluation%20Policy.pdf> (last accessed June 12, 2015); Agenda for August 5, 2011 State Board Meeting, Agenda Item III.M (concerning amendments to teacher evaluation rule), *available at* <http://state.tn.us/sbe/2011Augustpdfs/III%20M%20Educator%20Evaluation%20Policy%20&%20Rule%20-%20Non%20Public%20Schools.pdf> (last accessed June 12, 2015); Agenda for November 4, 2011 State Board Meeting, Agenda Item IV.G (concerning amendments to teacher evaluation policy in response to “widespread feedback from directors, principals, and teachers throughout the state”), *available at* <http://state.tn.us/sbe/2011Novemberpdfs/IV%20G%20Teacher%20&%20Principal%20Evaluation%20Policy.pdf> (last accessed June 10, 2015).

¹³ The court has identified publicly available versions of the policy that purport to have been revised on November 4, 2011 (the earliest post-Act revision located by the court), June 22, 2012, August 13, 2012, April 19, 2013, October 25, 2013 (the version attached to the Complaint), and January 30, 2015 (the most recent revision located by the court).

- Purpose: “The primary purpose of annual teacher and principal evaluations is to identify and support instruction that will lead to high levels of student achievement.”
- Use: “Evaluations will be used to inform human capital decisions, including, but not limited to[,] individual and group professional development plans, hiring, assessment, and promotion, tenure and dismissal, and compensation.”
- 50%/50% Split: In compliance with the then-existing statutory percentage requirements, the policy indicated that 50% of each teacher’s evaluation would be comprised of student achievement data, including 35% based on “student growth measures” and 15% based on other measures of student achievement, and that the remaining 50% would be based on a “qualitative appraisal instrument.”
- The 35% “Student Growth Component”: For the 35% individual student growth measure, the policy stated that the student growth measure would be the TVAAS for teachers with individual value-added scores. The policy stated that, for “teachers, librarians, counselors and other groups of educators who do not have individual TVAAS scores,” school districts “will choose from a list of options that have been shown capable of measuring student growth” and that have been approved by the Tennessee Department of Education before each school year. The policy stated that the State Board would “continually monitor and revise the list of options under this category based on increasing availability of high-quality measures of performance,” and that it “will work to develop valid and reliable student growth measures for those areas that do not currently have them.” However, “in lieu of the availability of growth measures for all educators without individual TVAAS scores, *school-level value-added scores will be the standard student growth measure while other growth measures are in development.*” (emphasis added). Although the policy recited that evaluators would choose from a list of alternative metrics for certain non-tested subjects, the policy did not actually include any alternative metrics. Thus, based on the face of the policy, teachers of non-tested subjects necessarily received a school-wide TVAAS composite score that constituted 25% of their evaluation.
- 15% Student Growth Component: For the 15% individual student growth measure premised on “other measures of student achievement,” the policy provided a list of nine different metrics from which evaluators and the evaluated teachers could choose, including the school-wide TVAAS composite and various other data points. For example, a “computer technology” teacher or a “world languages teacher” could be evaluated based on TCAP results, the school-wide TVAAS score, the SAT or ACT

“suite of assessments,” “National/State ‘off-the-shelf’ assessments based on criteria developed by the TDOE,” the AP/IB/NIC suites of assessment,” post-secondary matriculation and placement rates, students’ success in advanced coursework, or 9th-grade retention and promotion rates.¹⁴

The State Board appears to have revised the policy on June 22, 2012, August 13, 2012, and April 19, 2013 without changing the basic features set forth in the November 4, 2011 version, except (in relevant part) to tinker with some of the available options for the 15% “other measures of student achievement” option.¹⁵

At some point after April 19, 2013 (and consistent with the Act’s initial directives over three years earlier), the State Board finally approved three additional alternative metrics for certain teachers of non-tested subjects and amended Policy 5.201 accordingly.¹⁶ At least as of the October 25, 2013 revision to Policy 5.201, the State Board revised the policy to include three alternative options to the TVAAS as to non-tested subjects, including a K-2 assessment, a “Fine Arts Portfolio Model,” and a “World Languages Portfolio Model.” The policy also included further revisions to the list of “other measures” that would account for 15% of each teacher’s evaluation. Although the plaintiffs do not raise the issue, it appears that this version of the policy failed to reflect at least two aspects of the intervening April 2013 statutory amendments to the

¹⁴ <http://www.tn.gov/sbe/Policies/5.201%20Teacher%20and%20Principal%20Evaluation%20Policy%20-%20Update%202011.pdf> (revised November 4, 2011) (last accessed June 12, 2015).

¹⁵ See http://www.tn.gov/sbe/Policies/5.201_Teacher_and_Principal_Evaluation%20Policy.pdf (revised June 22, 2012) (last accessed June 12, 2015); http://www.tn.gov/sbe/Policies/5.201_Teacher_and_Principal_Evaluation_Policy_Revised_9-28-2012.pdf (revised August 13, 2012) (last accessed June 12, 2015); http://www.state.tn.us/sbe/Policies/5.201_Teacher_and_Principal_Evaluation_Policy_revised_4-19-2013.pdf (revised April 19, 2013) (last accessed June 12, 2013).

¹⁶ See Complaint, Ex. A (Policy 5.201, revised October 25, 2013).

Act: first, the policy did not reflect the amendment reducing the “student achievement data” component from 50% to 40% for teachers without access to individualized data (*see* Policy 5.201(1)(a)); second, it failed to reflect the amendment permitting teachers (rather than evaluators) to choose the 15% “other measures” metric (*see* Policy 5.201(1)(b)(1)).

The most current version of the policy purports to have been revised as of January 30, 2015, and it appears to be effective for the 2014-15 school year forward. *See* http://www.tn.gov/sbe/Policies/5.201_TeacherandPrincipalEvaluationPolicy_1-30-2015.pdf (last accessed June 12, 2015). In this version of the policy, the State Board has approved a fourth alternative to the TVAAS: a Physical Education Student Growth Model. As to teachers of non-tested subjects who are not covered by the four additional alternatives, the policy continues to require that those teachers be evaluated, in part, based on the school-wide TVAAS composite score. Belatedly reflecting one of the April 2013 changes to the Act, the policy also indicates that, where there is a disagreement between a teacher and the evaluator about what metric to use for the 15% “other measure of student achievement” component, the teacher gets to choose. Again, however, the policy continues to fail to acknowledge that, for teachers without access to individualized growth data, only 40% of their evaluation — not 50% — may be based on student achievement data.

III. District-Level Policies Implementing the Act and Policy 5.201

The Complaint alleges that districts have taken a variety of approaches to implementing the Act and the State Board Policy 5.201, although one feature remains constant: in every district, at least some teachers of non-tested subjects are evaluated based on school-wide TVAAS data. The Complaint contains both general statewide allegations and allegations specific to the MNBPE, which oversees the Nashville District, and the ACSBE, which oversees the Anderson

District. Both districts utilize separate policies to implement the Act and State Board Policy 5.201, which are attached to the Complaint and which the court will refer to as to the Nashville Policy and the Anderson Policy.

According to the Complaint, school districts have adopted different approaches as they relate to teachers of non-tested subjects. Some districts have maximized the use of TVAAS data to evaluate teachers of non-tested subjects: for the 2011-12 and 2012-2013 school years, those districts based 50% of each teacher's evaluation on the school-wide TVAAS score and 50% on personal observation; and for the 2013-2014 school year, they based 40% of the evaluation on the TVAAS score (the maximum allowable) and 60% on personal observation. The plaintiffs allege that some of those districts, including the Nashville District through the Nashville Policy, have continued to utilize the TVAAS school-wide composite *regardless* of whether an individualized alternative to the TVAAS has been made available under Policy 5.201. By contrast, on the opposite extreme, other districts may have minimized (at least to some degree) the use of the TVAAS scores to evaluate teachers of non-tested subjects. Those districts utilize the TVAAS school-wide score only where an alternative metric is not available and, even then, only granting it the minimum weight allowable under Policy 5.201 and the Act (namely, 35% of the evaluation for the 2011-12 and 2012-13 school years and 25% of the evaluation for the 2013-14 school year forward). The operative point is that, in school districts across the state, including the Nashville District under the Nashville Policy and the Anderson District under the Anderson Policy, school-wide TVAAS composite scores now count for at least 25% of many teachers' individual performance evaluations.

More broadly, the plaintiffs allege that (1) for the 2011-12 and 2012-13 school years, more than half of Tennessee's teachers received a performance evaluation based between 35%

and 50% on school-wide TVAAS scores, and (2) for the 2013-14 school year, more than half of Tennessee's teachers received a performance evaluation based between 25% and 40% on school-wide TVAAS scores. The plaintiffs also allege that, even where alternative performance measures are available for the "student achievement" component (such as the K-2 Assessment or the World Languages Assessment), few school districts have utilized those alternative methods instead of the school-wide TVAAS composite. The plaintiffs allege that the same holds true for the 15% "other achievement" component, which often consists of a school-wide TVAAS score.¹⁷ Many teachers are therefore being evaluated based, in substantial part, on the test scores of students that they did not teach and in subjects that they do not instruct.

IV. Consequences of the Performance Rating

The Complaint alleges that a teacher's evaluation rating can have negative consequences for a teacher's compensation and other conditions of employment.

Teachers who receive a "1" or a "2" rating may be suspended for "inefficiency." *See* Tenn. Code Ann. §§ 49-5-501(6) and 49-5-511(a)(2). Teachers who have attained tenure after July 1, 2011 can be returned to probationary status if their performance rating in two consecutive years is a "2" or "1" – thereby causing them to lose protection from discharge, absent cause. *Id.* § 49-5-504(e)-(f). Furthermore, in fourteen of Tennessee's school districts, teachers are

¹⁷ As the court understands current State Board policy, this last point cannot – or at least should not – be valid for the 2014-15 year forward for many, if not all, teachers of non-tested subjects. Under the current policy, the evaluated entity (the teacher) has the right to choose what metric to use for this 15% component. The TVAAS school-wide score is just one of multiple options available under Policy 5.201. Therefore, a teacher evaluation that includes a 15% "other measures" component consisting of the TVAAS school-wide score must reflect that teacher's *choice* to utilize the TVAAS school-wide score, rather than one of the other options (such as the school's graduation rate rating).

eligible to receive a performance bonus under the Teacher Incentive Fund Grant (“TIF”) program, provided that they meet district-specific eligibility criteria. A teacher’s eligibility for a TIF grant can be negatively impacted by a performance rating that may include a TVAAS measure of performance.

V. The Plaintiffs and the Defendants

The plaintiffs in this case include two individual teachers (Theresa Wagner and Jennifer Brauener), the Metropolitan Nashville Education Association (“MNEA”) (a voluntary association of Nashville-area teachers and other professionals), the Anderson County Education Association (“ACEA”) (the Anderson County equivalent to the MNEA), and the Tennessee Education Association (“TEA”) (a statewide professional employees’ organization under Tenn. Code Ann. § 49-5-602(9)). For purposes of the instant motions, the court will assume that the plaintiffs collectively cover the interests of all potential teachers impacted by the laws and policies at issue. The defendants are Tennessee Governor Bill Haslam, Tennessee Commissioner of Education Candace McQueen, nine individuals who compose the State Board, the MNPBE, and the ACSBE.

VI. The Individual Plaintiffs’ Allegations

The two individual plaintiffs allege sympathetic facts that, if true, demonstrate how the use of school-wide TVAAS scores can have seemingly unfair results for teachers of non-tested subjects.

1. Theresa Wagner

Plaintiff Theresa Wagner teaches physical education to grades 5-8 at Gra-Mar Middle Prep in Nashville, Tennessee. She has taught physical education for 26 years, including seventeen years of teaching within the Nashville District and eleven years of teaching at Gra-Mar

Middle Prep. She serves as the school's Healthy Team Coordinator (for which she received a "Healthy Schools Champion" award in 2013-14), she serves as the Lead Teacher for the Teacher Incentive Fund, she manages the school's website, and she previously served on the district's Strategic Compensation Committee. She consistently receives positive performance reviews and is well-regarded by her colleagues and the school community. She is licensed to teach Physical Education for grades K-12 and has completed coursework for teaching Adapted Physical Education to students with disabilities.

The state-mandated curriculum that Ms. Wagner must teach is unrelated to the subject matter areas covered by the TVAAS. Ms. Wagner must teach specific motor skills such as running, dribbling, and balance skills, understanding and applying offensive and defensive strategies in activities such as basketball and tennis, and understanding the importance of maintaining physical activity.

For the 2011-12 and 2012-13 years, Ms. Wagner received overall performance ratings of "5" in both years, based on both (a) her observation results, and (b) the school-level TVAAS estimate of "5" in both years. During these years, she received the maximum TIF performance bonus each year (\$2000 and \$1000, respectively). However, in the 2013-14 school year, the school-wide TVAAS estimate dropped from "5" to "1." Because the MNBE utilized the TVAAS estimate for 40% of Ms. Wagner's evaluation for the 2013-14 school year, this precipitous drop in the TVAAS estimate dragged Ms. Wagner's overall performance rating down along with it: although her observation results remained high (presumably in line with her previous "5" ratings), her overall performance rating dropped to a "3." This rendered her ineligible for a TIF performance bonus and allegedly hurt Ms. Wagner's reputation, lowered her morale, and caused her emotional distress. The school-wide TVAAS score of "1" for that year

reflected student performance exclusively on standardized testing of subjects other than physical education and included the performance (in those other subjects) of students that Ms. Wagner did not teach at all that year.

2. Jennifer Braeuner

Plaintiff Jennifer Braeuner teaches Visual Arts to students in grades 6-8, including one special education class, at Norris Middle School in Norris, Tennessee, which is part of the Anderson District. She has taught there for six years. She has a Master's Degree in Art Education and she is licensed to teach Visual Arts in grades K-12. She sponsors the school art club, runs a weekly school-wide design contest to engage students in art outside the classroom, serves as a track and field coach and a substitute member of her School Leadership Team, and previously served as a mentor for graduating high school students. Her students' artwork has received public recognition, including one student whose artwork was selected for a special exhibition at the Knoxville Museum of Art and another student whose artwork won a district-wide poster contest. She has consistently received positive performance reviews based on actual observations of her teaching. She is recognized by colleagues and the school community as an effective teacher.

The middle school art curriculum that Ms. Braeuner is required to teach includes the understanding and use of different art media and the understanding of the elements of art and design. The TCAP examination (applicable to the grade levels that Ms. Braeuner teaches), do not test the standards that she is licensed, employed, and required to teach. For example, the TCAP Sixth Grade reading examination tests for the correct use of nouns, verbs, adjectives, prepositional phrases, synonyms, and antonyms and the selection of appropriate thesis

statements, supporting statements, and concluding sentences for a writing sample, as well as other aspects of the required 6th-grade English curriculum.

For the 2011-12, 2012-13, and 2013-14 school years, Ms. Braeuner's performance evaluation was based in part on Norris Middle School's schoolwide composite score, which was based entirely on standardized testing in subjects that did not include the visual arts and based in part on the performance of students at Norris Middle whom she did not teach. She expects that the situation will be the same for the 2014-15 school year. For the school years 2011-12 and 2012-13, she received an overall "5" performance rating because of high marks on personal observation results and because the school had a TVAAS school-wide estimate of "5." However, for the 2013-14 school year, the school-wide composite rating fell to "1," thereby dragging her overall performance rating down with it: although her observational performance rating remained high, she received an overall "3" rating. Because of this drop in her overall performance rating, she became ineligible for consideration for tenure (at least for the time being), suffered harm to her professional reputation, and experienced diminished morale and emotional distress.

The TEA, MNEA, and ACEA allege that they represent many teachers with similar histories.

VII. The Plaintiffs' Constitutional Challenges

In the Complaint, the plaintiffs challenge the constitutionality of Policy 5.201, the Nashville Policy, and the Anderson Policy. They allege that, both facially and as applied, all three policies violate the Due Process Clause and the Equal Protection Clause of the Fourteenth Amendment to the United States Constitution, to the extent that they require teachers of non-tested subjects to be evaluated using school-wide TVAAS scores.

The plaintiffs claim that they are challenging only the policies, not the Act itself. They claim it is arbitrary, capricious, and irrational to require that teachers “be evaluated substantially on the basis of student standardized test scores unrelated to the courses they teach” and that the policies utilizing that approach violate due process. (*See* Compl. ¶ 60.) They also allege that the policies violate the Equal Protection Clause because there is no rational reason to assign individual TVAAS scores to teachers of tested subjects while assigning school-wide TVAAS scores to teachers of non-tested subjects. The plaintiffs have sued the individual defendants in their official capacities, along with the MNBPE and the ACSBE. The plaintiffs ask the court to (1) declare that the policies are unconstitutional, to the extent that they require teachers of non-tested subjects to be evaluated using TVAAS data, (2) enjoin the defendants from implementing or enforcing those provisions of the policies, and (3) award the plaintiffs their fees.

The State Defendants (Governor Haslam, Commissioner McQueen, and the State Board members) have filed a Motion to Dismiss under Rule 12(b)(6). They contend that (1) the plaintiffs’ facial challenges to Policy 5.201 do not meet the standard set forth in *United States v. Salerno*, 481 U.S. 739, 745 (1987); and (2) there is a rational basis for Policy 5.201. Their motion is supported by a Memorandum of Law. (Docket No. 31, Attach. No. 1.)

Defendant MNBPE has filed a separate Motion to Dismiss under Rule 12(b)(6). The MNBPE asserts three arguments. First, it contends that it cannot be held liable for implementing the challenged provisions of the Nashville Policy because those provisions were mandated by the State Policy or the Act itself, thereby leaving the MNBPE no choice but to utilize TVAAS scores for teachers of non-tested subjects. Second, the MNBPE argues that, even if it had some choice in the matter, the Nashville Policy requiring the use of the TVAAS in all performance evaluations is constitutional. Third, the MNBPE argues that it is not an entity subject to suit.

**RULE 12 STANDARD FOR MOTIONS TO DISMISS RATIONAL BASIS
CONSTITUTIONAL CHALLENGES**

The standard for an equal protection claim is set forth in *Midkiff v. Adams Cnty. Reg'l Water Dist.*, 409 F.3d 758 (6th Cir. 2005):

To state a claim under the Equal Protection Clause, a § 1983 plaintiff must allege that a state actor intentionally discriminated against the plaintiff because of membership in a protected class or burdened a fundamental right. If a protected class or fundamental right is involved, this Court must apply strict scrutiny, but where no suspect class or fundamental right is implicated, this Court must apply rational basis review. Under rational basis review, the governmental policy at issue will be upheld as long as there is a rational relationship between the disparity of treatment and some legitimate government purpose. Under rational basis review, a plaintiff faces a severe burden and must negate all possible rational justifications for the distinction.

Id. at 770 (internal quotations and citations omitted). “There is a rational basis for a governmental classification if the distinctions made by that classification ‘have some relevance to the purpose for which the classification is made.’” *Darks v. Cincinnati*, 745 F.2d 1040, 1042-43 (6th Cir. 1984) (quoting *Marshall v. United States*, 414 U.S. 417, 422 (1974)). The standard for a claim under the Due Process Clause is substantially the same. *See Midkiff*, 409 F.3d at 769 (stating that substantive due process requires only that a policy bear a “rational relationship” to a “legitimate government interest.”) On rational basis review, the policy being challenged bears “a strong presumption of validity.” *F.C.C. v. Beach Commcn’s, Inc*, 508 U.S. 307, 315 (1993). Moreover, because policymakers are not required to articulate their justification for a policy, “it is entirely irrelevant for constitutional purposes whether the conceived reason for the challenged distinction actually motivated the legislature.” *Id.* at 316. “[A] legislative choice is not subject to courtroom fact-finding and may be based on rational speculation unsupported by evidence or empirical data.” *Id.*

The Sixth Circuit recently explained the relationship between the Fourteenth Amendment and constitutional challenges to state laws in more detail:

A first requirement of any law, whether under the Due Process Clause or Equal Protection Clause, is that it rationally advance a legitimate government policy. Two words (“judicial restraint”) and one principle (trust in the people that “even improvident decisions will eventually be rectified by the democratic process”) tell us all we need to know about the light touch judges should use in reviewing laws under the standard. So long as judges can conceive of some plausible reason for the law – *any* plausible reason, even one that did not motivate the legislators who enacted it – the law must stand, no matter how unfair, unjust, or unwise the judges may consider it as citizens.

. . . The signature feature of rational basis review is that governments will not be placed in the dock for doing too much or for doing too little in addressing a policy question. . . . [R]ational basis review does not permit courts to invalidate laws every time a new and allegedly better way of addressing a policy emerges, even a better way supported by evidence and . . . by judicial factfinding. If legislative choices may rest on “rational speculation unsupported by evidence or empirical data,” it is hard to see the point of premising a ruling on unconstitutionality on factual findings made by one unelected federal judge that favor a different policy. Rational basis review does not empower federal courts to “subject” legislative line-drawing to “courtroom” factfinding designed to show that legislatures have done too much or too little.

DeBoer v. Snyder, 772 F.3d 388, 404-405 (6th Cir. 2014) (citations omitted).

ANALYSIS

I. The Rational Basis Standard

Teachers and other members of the public may have a keen interest in the outcome of this litigation. Given that many non-lawyers may read this opinion, it may be helpful to explain in layman’s terms just how limited the court’s role is here.

The United States Constitution says nothing about public education. In some spheres, the federal government can pass legislation concerning public education (Title IX, for example). Where the federal government acts, the Supremacy Clause of the Constitution generally mandates that the federal law trumps a conflicting state law or policy, if there is one. The Equal

Protection Clause of the Fourteenth Amendment to the United States Constitution also protects certain specific groups from discrimination. These protected groups include, among others, women, racial minorities, and ethnic minorities, and these groups are generally referred to in constitutional law as “suspect classes.” The Fourteenth Amendment also protects certain “fundamental rights,” such as the right to freedom of religion. When a law discriminates against a suspect class or places a burden on a fundamental right (for instance, a policy that requires females in a school to be treated differently than males in the school for disciplinary purposes), the government must show both that it has discriminated in pursuit of an important objective and that the policy is appropriately tailored to achieve that end. Thus, when a state policy impacts a suspect class or a fundamental right of all U.S. citizens, a federal court must scrutinize the law to determine whether the means are appropriate to achieve an important end. Similarly, the Due Process Clause of the Fourteenth Amendment also ensures that the government may not deprive citizens of “life, liberty, and property” without “due process” of law. This clause protects citizens from the deprivation of certain interests, such as a legitimate property claim of entitlement to a government service, and even in those instances it protects citizens only from the deprivation of those interests without an appropriate process for doing so. The clause also protects citizens from the deprivation of particular constitutional rights and from actions that shock the conscience.

Unless one of these limited exceptions is implicated, the United States Constitution grants the states nearly unfettered discretion to regulate public education. In most situations, as is the case here, the Constitution requires only that the state or local policy be “rationally related” to a “legitimate governmental objective.” This is the lowest form of constitutional scrutiny that federal courts apply, and it is exceedingly lenient in favor of policymakers.

The term “rational” has a special meaning in constitutional law. It does not mean that a law must be reasonable, sensible, well-founded, effective, or supported by evidence. In fact, in many contexts (including public education), state legislatures and policymakers can make “rational” decisions that, when implemented, are unreasonable, counter-productive to a stated goal, or contrary to all past experience and evidence. When a federal court conducts a rational basis review, the review is limited to determining only whether there is a *conceivable* rational *relationship* between the policy and a legitimate governmental objective. The rational relationship need not even be articulated by the policymakers defending the policy choice, and there may be no actual proof supporting that decision (in fact, there could even be proof demonstrating that the policy does not, in fact, achieve the desired result).

In sum, where rational basis review applies – as is the case here – the U.S. Constitution allows state legislators and policymakers to make both excellent decisions and terrible decisions, provided that the decisions are based on some conceivable modicum of rationality at the time of their passage or application in practice. The U.S. Constitution does not permit a federal court to evaluate or rule upon the wisdom of these decisions, even where the policy may be unfair, misguided, or counter-productive. Thus, when a federal court finds that a policy is “rationally related to a legitimate government objective,” the court is not endorsing the policy, finding that it is empirically supported, or concluding that it is a wise idea. The court is merely ruling that the U.S. Constitution *does not forbid* a state or locality from adopting or applying that policy.

It is within this context that the court must evaluate the state and district policies challenged by the plaintiffs.

II. The Act and Policy 5.201

The plaintiffs do not dispute that it is legitimate for the government to implement teacher evaluation policies to “identify and support instruction that will lead to high levels of student achievement,” as set forth in Policy 5.201. However, the gravamen of the plaintiffs’ Complaint is that it is irrational to utilize school-wide TVAAS scores as a component of an individual teacher’s individual performance evaluation, because the TVAAS does not actually measure the individual performance of those teachers.

A. Introduction

Complicating the court’s analysis somewhat is the fact that the plaintiffs have waited for five and one-half years to assert these constitutional challenges. The policies being challenged have been in effect, in one form or another, since the 2011-12 school year. In the intervening years, the Act, Policy 5.201, and the various district policies have changed materially multiple times. The result is a hodgepodge of variations, across time and geography, in the manner in which Tennessee teachers have been evaluated since the Act’s passage. Furthermore, neither the Complaint nor the plaintiffs’ briefing parse out with specificity which aspects of the policies being challenged are (a) required of the State Defendants under the Act at any given point in time or (b) required of the districts by Policy 5.201 at any given point in time. This has led to some disagreement in the parties’ briefing about whether the plaintiffs are actually challenging the Act itself, challenging each policy “on its face,” or challenging each policy “as applied.”

As an initial matter, it is difficult to see how the plaintiffs are not challenging the Act itself, even as they disclaim that they are attacking its constitutionality. The Act indicates in multiple places that VAAs should be used for at least 35% (and later at least 25%) of each teacher’s evaluation, and the Act appears to contemplate that developing VAAs for many subject matter areas would not occur all at once, but rather incrementally with input from the TEAC and

the State Board. The implication is that, from the outset, the Act contemplated that some teachers would receive VAA ratings that were not individualized, but only until alternative individualized metrics could be validated and implemented.

Nevertheless, for the sake of argument, the court will assume that the State Board had some discretion in identifying alternative metrics to the TVAAS that would have covered all non-tested subject matter areas. The court will also treat the constitutional challenges to Policy 5.201 as “as applied” challenges subject to rational basis review, which is a marginally stricter standard than a facial challenge under the *Salerno* standard.

B. Rational Basis Analysis of Policy 5.201

The school-wide TVAAS score reflects the performance of all students at the school on certain tested subjects. It was (and is) rational for the State Board to utilize school-wide TVAAS scores, which measure those students’ performance on those subjects, as a component of individual teacher evaluations for teachers of non-tested subjects.

There are several reasons why, with respect to teachers of non-tested subjects, the State Board could rationally believe that a school-wide score provides some measure (albeit a crude one) of evaluating an individual teacher’s performance. Policymakers could believe that a teacher has an effect – positive or negative – on a student’s performance on statewide assessments in other subjects. They could also believe that the quality of a teacher’s instruction impacts the school as a whole and that teachers should be encouraged to invest in raising the bar for the entire school. A teacher with that incentive might be more likely to participate in school-wide activities or to encourage her students to do so, she might produce students who are better behaved, and she might create students who become better learners in other classes. For example, students of an effective physical education teacher might be healthier and more

attentive in other classes and during exams, and the teacher might set an example for the whole school by serving as a mentor to other students or by running a school-wide scholastic competition. Students who perform well in music class might demonstrate better proficiency on mathematics or other subjects requiring specialized focus.¹⁸ Furthermore, utilizing a school-

¹⁸ Even a cursory search reveals empirical support, both in the scholarly literature and reports from federal government agencies and task forces, for the notion that performance in a non-core subject can improve standardized tests scores in other subjects. See, e.g., Arnoud Cabanac, *et al.*, “Music and Academic Performance,” *Behavioural Brain Research*, Vol. 256, at pp. 257-260 (Nov. 1, 2013), available at <http://www.sciencedirect.com/science/article/pii/S0166432813005093> (last accessed June 12, 2015) (confirming hypothesis that students studying music have better grades in all subjects); Marie Forgeard, *et al.*, “Practicing a Musical Instrument in Childhood is Associated with Enhanced Verbal Ability and Nonverbal Reasoning,” *PLoS ONE* 3(10): e3566 (Oct. 29, 2008), available at <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0003566> (last accessed June 12, 2015) (finding at least a correlational relationship between fine instrumental music training and enhanced auditory discrimination, fine motor skills, vocabulary, and nonverbal reasoning); Centers for Disease Control and Prevention, “The Association Between School Based Physical Activity, Including Physical Education, and Academic Performance.” Atlanta, GA: U.S. Department of Health and Human Services (revised July 2010), available at http://www.cdc.gov/HealthyYouth/health_and_academics/pdf/pa-pe_paper.pdf (last accessed June 12, 2015) (stating, *inter alia*, that “[t]here is substantial evidence that physical activity can help improve academic achievement, including grades and standardized test scores,” and that “physical activity can have an impact on cognitive skills and attitudes and academic behavior, all of which are important components of improved academic performance”); President’s Committee on the Arts and the Humanities, “Reinvesting in Arts Education: Winning America’s Future Through Creative Schools” (May 2011), available at http://www.pcah.gov/sites/default/files/photos/PCAH_Reinvesting_4web.pdf (last accessed June 12, 2015) (stating that “[a]ll of the research points to the success of schools that are ‘arts-rich’ – in which students who may have fallen by the wayside find themselves re-engaged in learning when their enthusiasm for film, design, theater or even hip-hop is tapped into by their teachers. More advanced students also reap rewards in this environment, demonstrating accelerated learning and sustained levels of motivation.”); Steven N. Kelly, Ph.D, “A Comparison of Cohort Data From 2007-2008 to 2010-11 Regarding Fine Arts-Related Instruction’s Influence on Academic Success,” Narrative Summary (Oct. 2012), available at <http://newsroom.pcsb.org/wp-content/uploads/2013/02/2010-2011-Cohort-Study.pdf> (last accessed June 12, 2015) (concluding that students participating in arts-related classes in Florida “1) stay in school; 2) have higher

wide score might also incentivize teachers to incorporate concepts from other tested subjects into their coursework. For example, a computer technology teacher could have the class design a website describing different events in U.S. History, or a physical education teacher could incorporate mathematical concepts into gym class, such as how geometry or physics plays a role in particular physical activities. The Constitution does not preclude policymakers from creating these incentives or from encouraging teachers to “teach to the test.”

Moreover, the use of organization-wide evaluation criteria is commonplace in many contexts. As the defendants point out, salespeople, attorneys, and managers often receive employment benefits based upon an organization’s collective success, such as increased sales or profits. Of course, it is likely that a certain salesperson (or teacher) contributed more or less to the organization’s success (or lack thereof). Even in that case, it would not be irrational for an organization to utilize the organization’s overall success both to incentivize employees to invest in the organization as a whole and as a loose proxy for each member’s performance. The same could hold true for teachers within a school. In a sense, policymakers could rationally reach the conclusion that “a rising tide lifts all boats” – or, by contrast, that a receding tide will sink them. Thus, it is rational to believe that utilizing a school-wide measure of performance provides some rough indication of the value added by an individual teacher within that school.

The plaintiffs also argue that it is unfair to evaluate teachers of tested subjects based on individualized scores while evaluating teachers of non-tested subjects based on school-wide scores. While that distinction may produce unfair results, such as for plaintiffs Wagner and Brauener in certain school years, it does not mean that the distinction is irrational. Policymakers

graduation rates; 3) perform higher in academic areas such as math, reading and writing; and 4) achieve higher scores on standard tests”).

can make policies that address a policy issue in part and leave further refinement for later. Here, in the absence of available alternative VAAs for certain subjects, the State Board has required that teachers be evaluated, in part, based on the TVAAS school-wide scores. The State Board has continued (albeit slowly) to develop and approve additional metrics to cover more subject matter areas. To be sure, until alternative metrics are available for particular teachers, the teacher of a tested subject may be left in a better position than the teacher of a non-tested subject (provided that the school-wide score is lower than the personal observation score of the teacher of the non-tested subject). But the U.S. Constitution does not require Tennessee or the State Board to choose between using individual TVAAS scores for all teachers or for none of them. Instead, it permits the legislature and the State Board to act incrementally, as they have, and to use a rational alternative for evaluating some teachers, where no “best” option is otherwise available.

The policy is also rational in context. First, at least for some school years, the policy indicates that some teachers of non-tested subjects should be tested based on individualized alternatives to the TVAAS. Second, even where the TVAAS school-wide composite was utilized because no alternative had been approved, the policy did not mandate just how much weight should be placed on the teacher’s overall performance evaluation, only that it “inform” (but not necessarily dictate) certain terms and conditions of teachers’ employment. Third, since its inception, the policy has permitted school districts or teachers to utilize metrics other than the TVAAS school-wide composite for 15% of each teacher’s evaluation. In other words, the policymakers built a number of qualifications and alternatives into Policy 5.201 that can alleviate the potential effect, if any, that the TVAAS school-wide score for a particular school has on teachers of non-tested subjects within that school.

By contrast, one can conceive of performance metrics that would be truly irrational, such as basing a Tennessee teacher's evaluation on the test scores of students in Arizona, whether the Nashville Sounds baseball team had a winning season that school year, or the State of Tennessee's economy on evaluation day. It is inconceivable that a Tennessee teacher's "value added" to a student's performance would bear any relationship to those metrics. By contrast, it is rational to believe that a teacher can impact the school-wide performance of both her own students and other students at that school by being an effective teacher and by improving the overall educational environment at the school. Certainly, the relationship between the actual value of a teacher's performance and the entire student body's performance in other subjects is attenuated, at best, and it is questionable whether it would be helpful or counter-productive to a particular class (say, a physical education class) to adopt an incentive that may encourage a teacher to incorporate other subjects (on which the teacher is not licensed) into a course. But it is not the province of this court to make value judgments about the wisdom of the State Board's policy choices based on the directives set forth in the Act. While it is a less-than-perfect method of evaluating individual teachers, using the school-wide TVAAS data is a "rational" policy choice that the federal Constitution does not forbid.

A federal court in Florida recently reached the same conclusion in response to a challenge to a Florida law that is substantially similar to the Act and the policies at issue here. *See Cook v. Stewart*, 28 F. Supp. 3d 1207 (N.D. Fla. 2014); *see also Kim Cook, et al. v. Pamela Stewart, et al.*, Case No. 1:13-cv-72-MW-GRJ (N.D. Fla. Apr. 22, 2014) (Docket No. 111) ("*Cook* Motion to Dismiss Order"). In *Cook*, the district court expressed concerns with "serious flaws . . . with the system" utilized in Florida, but concluded that the law at issue "survives rational review even

if it seems unwise or if the rationale for it seems tenuous.” *Cook* Motion to Dismiss Order at p. 17. This court reaches essentially the same conclusion.

Although they are not necessary to the court’s rational basis review, two final points are warranted.

First, it is not lost on the court that the policies at issue did not arise out of thin air. The Tennessee legislature decided to utilize VAAs as a partial measure of teacher performance based, according to the statute, on research studies relating to their effectiveness. The Act created the TEAC, which is composed of a variety of individuals who are invested in Tennessee’s education system and who are charged with fashioning teacher evaluation policies for teachers across the state. The TEAC makes recommendations to the State Board, which itself is composed of a diverse group of individuals who are directly involved in Tennessee’s public education system. Through appointments and required consultation under the Act, both the Governor and the legislature have an influence on this decision-making process at multiple levels. If this court were to declare the policies “irrational,” it would be tantamount to one unelected federal judge declaring that all of these policymakers, who undoubtedly have devoted significant time and effort to fashioning feasible teacher evaluation policies, could not rationally have determined that approximately 25% of a teacher’s evaluation should be based on a school-wide performance score. Although there are cases in which the Constitution compels a federal court to override policymakers, this case is not one of them, and the circumstances indicate that the legislative and executive branches are well-positioned to sort out problems with the Act and its implementation.

This first point dovetails with the second: beginning soon after its passage and continuing through the current legislative session, the Tennessee legislature has continued to amend the Act periodically, often in ways that address the concerns raised by the plaintiffs, and so has the State

Board to some degree. For example, the legislature has given teachers the option of electing the “other measure” of student growth comprising 15% of the performance rating, it has now reduced the proportion of the evaluation attributable to a TVAAS score twice—from 35% to 25% to 10-15%— and it has clarified, if it were not clear before, that the Act does not require school districts to utilize student achievement data based on state assessments as the sole factor in employment decisions. The same holds true for the State Board, which continues to revise Policy 5.201 as it relates to alternative performance metrics. This is the political process at work, refining and potentially improving a law based on experience and feedback – as it should be.

In sum, the court finds that State Policy 5.201 (and the associated dictates in the Act) are rationally related to a legitimate government objective. The plaintiffs’ facial challenge and as applied challenges to State Policy 5.201 therefore fail.

III. The MNBPE’s Motion to Dismiss

A. Lack of Discretion Defense

When a municipality merely complies with state mandates that afford no discretion, such that they “act as an arm of the State,” the municipality cannot be held legally responsible under § 1983 for a constitutional violation resulting from compliance with that mandate. *Brotherton v. Cleveland*, 173 F.3d 552, 566 (6th Cir. 1999). The MNBPE argues that Policy 5.201 required – and continues to require – that the MNBPE utilize the TVAAS to evaluate teachers of non-tested subjects.

As the plaintiffs point out, the MNBPE’s position is only partially correct. Policy 5.201 states that districts should utilize alternatives to the school-wide TVAAS composite for teachers of K-2, the fine arts, world languages, and (as of the 2014-15 school year) physical education.

For whatever reason, the MNBPE (along with other districts) allegedly has chosen not to utilize these alternatives with respect to teachers of those particular subject matter areas. Because the MNBPE has discretion relative to these specific types of teachers, its decision not to utilize the alternative metrics is subject to constitutional scrutiny. As to all other teachers of non-tested subjects, the MNBPE has no discretion and cannot be held liable for utilizing the TVAAS as to them.

B. Whether the MNBPE is a Proper Defendant

The MNBPE argues that it is not subject to suit because, under the Metropolitan Charter, it was not granted the power to sue and be sued. *See Haines v. Metro. Gov't of Davidson Cnty.*, 32 F. Supp. 991, 994 (M.D. Tenn. 1998). The defendants argue that two post-*Haines* cases from Tennessee state courts indicate that a board of education can sue or be sued in certain contexts. *See S. Constructors, Inc. v. Loudon Cnty. Bd. of Educ.*, 58 S.W.3d 706, 715-718 (2001); *Byrn v. Metro. Bd. of Public Educ.*, 1991 WL 7806, at *4 (Tenn. Ct. App. 1991). Although this may be an issue of some complexity, the court finds no reason to construe *Southern Constructors* or *Byrn* as inconsistent with this court's reasoning in *Haines*. Both *Southern Constructors* and *Byrn* involve district-specific considerations related to the specific contract-related rights that Tennessee has conferred upon particular localities, not the considerations specific to the Metro Nashville Charter that this court scrutinized in *Haines*. Therefore, the MNBPE is not subject to suit, even with respect to the narrow areas in which it has had discretion under Policy 5.201.

Of course, this defect is only a nominal problem that is easily cured. The court could either construe the claims as asserted against Metro Nashville (the lawyers are the same) or the court could permit the plaintiffs to file an amended complaint that names Metro Nashville. Rather than prolong the inevitable, the court will construe the claims as asserted against Metro

Nashville itself, and the court will rule on the constitutionality of the relevant aspects of the Nashville Policy because the constitutional analysis is the same.

C. Rational Basis Analysis for Teachers of Subjects Covered by a State-Approved Alternative Individualized Performance Metric.

It seems very unfair to continue to utilize the TVAAS to evaluate teachers of non-tested subjects, despite the availability of an alternative individualized subject-matter specific test. Nevertheless, the decision to utilize the TVAAS school-wide score in that circumstance is not irrational.

For the reasons described in the previous section, it is rational to believe that the TVAAS provides some measure of teacher performance, even for teachers of non-tested subjects. The fact that an individualized test might provide a *better* measure of performance than the TVAAS school-wide composite does not mean that the Fourteenth Amendment requires it. Where rational basis review applies, the Constitution does not require policymakers to choose the best available option, only that policymakers choose an option that is rationally related to a legitimate governmental objective, which the MNBPE has done here.

IV. Additional Concerns

The policies at issue meet the minimal standard of rationality required by the U.S. Constitution. Nevertheless, there may have been significant problems with the pace and nature of the Act's implementation. Although the plaintiffs do not frame it this way, they may have a better argument that the State Board or the TEAC have only partially carried out their duties under the Act and that some local districts may be violating Policy 5.201 itself.

The Act created the TEAC and charged it with fashioning a policy for teacher evaluations, and it permitted state regulators to approve alternative measures for different subject

matter areas. At the time that the Act was passed, it may be that the legislature and the laws' supporters envisioned that, with input from the TEAC, the relevant agencies would expeditiously create and implement value-added measures of student achievement for many additional subject matter areas, in order to provide a tailored measure of each teacher's performance. They may have envisioned that school-wide TVAAS scores, while a crude measure of individual performance at best, might have served as only a temporary measure, acting only as a "bridge" until the TEAC and the State Board could find other, better measures of individualized performance to supplant the TVAAS for many (perhaps all) remaining subject matter areas.¹⁹

If the plaintiffs' allegations are true, that expectation has not come to pass: in the five and-one-half years since the Act's passage, the State Board has only approved a handful of metrics for evaluating teachers of non-tested subjects. Instead, the teachers of most non-tested subjects are left in the lurch, and it raises a question as to whether school-wide TVAAS scores bear a "strong relationship to the core curriculum for the applicable grade level and subject" taught by those teachers, as the statute appears to require. *See* § 603(c). Of course, there may be valid reasons why only four alternative metrics have been developed: perhaps no one has yet developed a valid and reliable value-added assessment for a computer technology teacher, for example. Whatever the reason, the issue merits continued attention by the legislature and the State Board. The fact that the legislature continues to reduce the impact that TVAAS scores have on certain teachers indicates that the political branches are indeed reacting to the types of criticisms leveled by the plaintiffs in this case.

¹⁹ Incidentally, it is troubling that the terms of Policy 5.201 have seemed to lag behind the enactment of certain amendments to the Act. For example, as far as the court can tell, the April 2013 reduction in the TVAAS percentage component for some teachers has never been incorporated into Policy 5.201.

Finally, although the plaintiffs do not raise the issue, there is a reasonable argument that State Board Policy 5.201 does *not* give local school districts discretion to utilize the TVAAS school-wide scores for teachers otherwise subject to approved alternative metrics. The policy states that school districts “will choose” from these four alternative metrics, not that school districts “may” or “can” choose from that list. Policy 5.201(1)(a) (emphasis added). This language could be construed as mandatory, not optional. If, as the plaintiffs allege, some districts are in fact refusing to apply these metrics in favor of the TVAAS school-wide composite (for example, by using the TVAAS school-wide score to evaluate a fine arts teacher), those districts may be violating Policy 5.201 itself. If that is the case, this would present an implementation problem for state courts to resolve, not a federal constitutional issue.

In sum, without endorsing the merits of the policies at issue, the court finds that Policy 5.201 and the Nashville Policy survive minimal constitutional scrutiny. The court will therefore grant the motions and dismiss all claims other than those asserted against the ACSBE.

CONCLUSION

For the reasons stated herein, the motion will be granted and the plaintiffs’ claims against all defendants other than the ACSBE will be dismissed with prejudice.

An appropriate order will enter.



ALET A. TRAUGER
United States District Judge