Notes

Teacher Evaluations Based on Student Testing: Missing an Opportunity for True Education Reform

Audrey R. Lynn*

I. INTRODUCTION: TEACHER EVALUATION METHODS 204

II. TEACHER EVALUATION POLICY IN THE STATES 206
- A. Race to the Top 206
- B. Controversy in the States About Teacher Evaluation Methods 207

III. CASE STUDY: FLORIDA 208
- A. The Florida Statute 209
- B. Delegation to Agencies for Rulemaking and History of Petition 210
- C. Rule and Petition Arguments 212
- D. Administrative Law Decision About Procedural Arguments 215
- E. Substantive Arguments 215
  - *1. The Rule Exceeded Grant of Rulemaking Authority and Enlarged, Modified, or Contravened* 217
  - *2. Vague and Lacks Adequate Standards* 223
  - *3. Arbitrary and Capricious* 225
- F. Summary of Substantive Argument Analysis 226

IV. PREVENTING LITIGATION 226

V. PUBLIC POLICY 229
- A. Efficacy of Tying Teacher Evaluations or Pay to Student Scores 229
- B. Work Environment for Teachers 231

VI. RECOMMENDATIONS 232

* I would like to thank the editors of the *Texas Journal on Civil Liberties & Civil Rights* for their help in preparing this Note for publication. I especially want to thank Professor Norma Cantú for her guidance in developing this Note.

I. INTRODUCTION: TEACHER EVALUATION METHODS

In the national movement to improve educational results, and especially in the current focus on test scores, the way teachers are evaluated is suggested by some as another way to improve public education. Many argue that the most influential school-based factor affecting student achievement is teacher quality.[1] Some see teacher evaluations as an important part of measuring teacher quality and are unimpressed with the current teacher evaluation systems, especially since evaluations often neglect measures of instruction quality and student learning in a direct and serious way.[2] An Economic Policy Institute study argued that "in practice, American public schools generally do a poor job of systematically developing and evaluating teachers."[3] Evaluations often consist of checklists based on short classroom visits, with some items on the checklist not directly addressing quality of instruction.[4] In addition to checklist items often being unsatisfactory, evaluators often give the same ratings to most of the teachers they evaluate and fail to discuss the evaluation results at a later date with the teachers.[5]

While many problems exist with current teacher evaluations, education experts offer a variety of potential solutions. Choices include more detailed standards that focus on indicators of instruction and planning ability, evaluations based on teacher portfolios as well as observations, and multiple evaluations by multiple evaluators.[6] Some of these systems focus more on improving instruction, rather than eliminating ineffective teachers.[7]

One method of teacher evaluation uses student test scores to score teacher performance. There are three test-based accountability program approaches used in the United States: status models, cohort-to-cohort change models, and value-added models.[8] No Child Left Behind (NCLB) combines two of these approaches to test-based accountability programs.[9] NCLB measures student scores of one group at one time against the state's annual measurable objective (a type of status model), while also comparing the change in student scores over time and increasing the state's annual measurable objective each year (a type of

---

[1] *E.g.*, JENNIFER KING RICE, TEACHER QUALITY: UNDERSTANDING THE EFFECTIVENESS OF TEACHER ATTRIBUTES 1 (2003).

[2] Thomas Toch, *Fixing Teacher Evaluation*, 66 EDUC. LEADERSHIP 32, 32 (Oct. 2008).

[3] EVA L. BAKER ET AL., ECOC. POLICY INST., PROBLEMS WITH THE USE OF STUDENT TEST SCORES TO EVALUATE TEACHERS 1 (2010), *available at* http://www.epi.org/publication/bp278/.

[4] Toch, *supra* note 2, at 32.

[5] *Id.*

[6] *Id.* at 32–35.

[7] *Id.* at 35.

[8] Daniel Koretz, *A Measured Approach*, 32 AM. EDUCATOR 2, 19 (2008), *available at* http://www.aft.org/pdfs/americaneducator/fall2008/koretz.pdf.

[9] *Id.*

cohort-to-cohort change model).[10] Some argue that tests are a "direct and simple" measurement tool.[11] Proponents of using student test scores as a part of measuring teacher performance assert that student test scores provide an effective measure of how well students are learning, especially when value-added data is used to help account for factors teachers cannot control.[12]

Using student test scores to evaluate teachers has recently taken the form of a value-added model, which attempts to "incorporate information on the value-added [sic] by individual teachers to the achievement of their students."[13] Value-added models generally compare a student's end-of-year test scores with test scores from the beginning of the school year and adjust the difference in scores by using other factors that may affect a student's score.[14] These factors are sometimes based on "student background or school-wide factors outside the teacher's control."[15] Value-added models are still being developed, and currently different formulations of these models are in use to measure the value a teacher adds each year.[16] Teachers and principals can gain valuable information about teachers' strengths and weaknesses from value-added measures of student test scores and can compare teacher performance with other teachers outside of their school.[17] Value-added models are generally believed to be a better way to measure performance than status models or cohort-to-cohort change models, but some argue that the information used by the model has limitations, making the model insufficient as a measure of teacher performance on its own.[18]

Some have suggested that student test scores should play a supporting role in evaluations, not a leading role.[19] However, many states have recently incorporated student test scores into a leading role in evaluations. These policy changes were partially triggered by Race to the Top. This Note proposes reduced reliance on student test scores when evaluating teachers, although student test scores—especially value-added models—could be used as a tool to improve teacher performance by highlighting areas that may need more work. First, this Note discusses recent policies in several U.S. states that have incorporated teacher evaluations of this kind; often, these policies result in conflicts in these

---

[10] *Id.*

[11] *Id.* at 21.

[12] *E.g.*, NAT'L COUNCIL ON TEACHER QUALITY, STATE OF THE STATES: TRENDS AND EARLY LESSONS ON TEACHER EVALUATION AND EFFECTIVENESS POLICIES 2–3 (2011), *available at* http://www.nctq.org/p/publications/docs/nctq_stateOfTheStates.pdf.

[13] STEVEN GLAZERMAN ET AL., BROOKINGS INST., BROWN CTR. ON EDUC. POLICY, EVALUATING TEACHERS: THE IMPORTANT ROLE OF VALUE-ADDED 2 (2010), *available at* http://www.brookings.edu/~/media/research/files/reports/2010/11/17%20evaluating%20teachers/1117_evaluating_teachers.

[14] *Id.*

[15] *Id.*

[16] Koretz, *supra* note 8, at 18.

[17] GLAZERMAN, *supra* note 13, at 4.

[18] Koretz, *supra* note 8, at 18–19.

[19] Toch, *supra* note 2, at 36.

states. Second, this Note focuses on Florida, a state that recently enacted such legislation as well as an agency rule; this section discusses the litigation and the rule it invalidated, and concludes that states should aim to avoid this kind of litigation. Third, this Note looks at ways to prevent the type of litigation that occurred in Florida from occurring in other states, including more collaboration with teacher groups in the policymaking process. Finally, this Note considers public policy reasons, including the existence of flaws in the value-added model system, to reduce reliance on student test scores when evaluating teachers.

II. TEACHER EVALUATION POLICY IN THE STATES

A. Race to the Top

In the past few years, more states are using teacher evaluations and teacher merit pay, both of which are at least partially based on student test scores. Race to the Top incorporates teacher evaluation reform into its selection criteria for determining which states will receive federal funds.[20] One criterion is the extent to which states "[d]esign and implement rigorous, transparent, and fair evaluation systems for teachers . . . that take into account data on student growth . . . as a significant factor . . . ."[21] Another criterion is whether states use these evaluations to "inform decisions" about teacher compensation, promotion, and retention.[22]

States are using this federal funding incentive to consider and sometimes implement merit pay and teacher evaluations based on student test scores.[23] Some teachers are reluctant to support the Race to the Top application, even though it would provide grants to their districts, because receipt of a grant will likely result in teacher evaluations based on student test scores.[24] For example, San Francisco teachers did not participate in the city's Race to the Top application process.[25] By contrast, politicians appear more likely to support participation in the Race to the Top program. For instance, Florida politicians embraced teacher evaluations based on test scores as a way of gaining Race to the

---

[20] U.S. DEP'T OF EDUC., RACE TO THE TOP PROGRAM EXECUTIVE SUMMARY 9 (2009), *available at* http://www2.ed.gov/programs/racetothetop/executive-summary.pdf.

[21] *Id.*

[22] *Id.*

[23] *E.g.*, Jeff Bishop, *There's Talk of Merit Pay for Teachers Through Race to the Top Program*, THE NEWNAN TIMES-HERALD (Georgia), Jan. 27, 2011, www.times-herald.com/Local/There-s-talk-of-merit-pay-for-teachers-through-Race-to-the-Top-program-1499986.

[24] John Fensterwald, *Skeptical Unions Pose Challenge to Districts' Race to the Top*, EDSOURCE (Sept. 20, 2012), www.edsource.org/today/2012/skeptical-unions-pose-challenge-to-districts-race-to-the-top/20320#.UJiBs4asTKd.

[25] *Id.*

Top funds. They introduced a bill basing teacher evaluations and compensation in part on student test scores.[26] The politicians described the bill's purpose as aligning with the state's Race to the Top application.[27] Former Florida Governor Jeb Bush supported merit pay, and current Governor Rick Scott signed a law bringing a merit pay system to Florida in March 2012.[28] Florida received a federal grant from Race to the Top to create a statewide teacher evaluation system.[29]

B. Controversy in the States about Teacher Evaluation Methods

The current focus on teacher evaluation and pay systems based on student test scores has gained traction as a way to improve education, but it has also received pushback in the form of litigation and political strife from affected parties, including teachers. Several states enacted evaluation systems based on student test scores or merit pay systems rewarding or penalizing teachers based on their students' scores, and these reform efforts have often met resistance.

For example, in California, a lawsuit was filed to enforce a California law requiring student progress to be included in evaluating teachers.[30] The judge in that case ordered both parties to create a new teacher evaluation system with a student progress measure as a part of the evaluation.[31] Teachers recently backed a proposed bill that would rewrite state rules about teacher evaluations, but the bill died in August 2012.[32] Some education advocates were pleased that the teachers' proposed bill failed, because a change in state rules could have weakened attempts to improve teacher quality with evaluations based upon student scores.[33] Of course, teachers took a different view.[34] The bill they supported provided for "a statewide uniform teacher evaluation system featuring more performance reviews, classroom observations, training of

---

[26] FLA. SENATE COMM. ON EDUC. PRE-K - 12, 2011 SUMMARY OF LEGISLATION PASSED 1, *available at* www.flsenate.gov/PublishedContent/Session/2011/Publications/BillSummary/PDF/ED0736.pdf.

[27] *Id.*

[28] Leslie Postal, *Gov. Scott Signs Teacher Merit-Pay Bill*, ORLANDO SENTINEL, March 24, 2011, http://articles.orlandosentinel.com/2011-03-24/news/os-florida-merit-pay-20110324_1_teacher-merit-pay-bill-judge-teacher-quality-national-teachers-union.

[29] *Id.*

[30] Stephen Sawchuk, *Critics Ask Calif. Courts to Change Teacher Policies*, EDUC. WEEK, June 6, 2012, at 21.

[31] Hillel Aron, *Concerns About Teacher Talks*, L.A. SCHOOL REPORT (Sept. 18, 2012), http://laschoolreport.com/how-far-away-are-utla-and-lausd-on-teacher-evaluations/.

[32] Teresa Watanabe & Michael J. Mishak, *California Teacher Evaluation Bill Abandoned by Lawmakers*, L.A. TIMES, Sept. 1, 2012, http://articles.latimes.com/2012/sep/01/local/la-me-teacher-eval-bill-20120901.

[33] *Id.*

[34] *Id.*

evaluators and public input into the review process."[35] Teachers stated that the features of their proposed evaluation system would have given clear guidelines to help teachers improve their instruction.[36] Implicit in the two sides' arguments was a difference in how teacher evaluations should be used. Teachers viewed evaluations as a tool for improvement, while school reform advocates and some parents viewed evaluations as a way to dismiss teachers who were not performing well enough.

In addition, a Louisiana teachers' association filed a lawsuit challenging a law that changes the teacher tenure system.[37] The law bases teacher tenure on student performance, puts teacher tenure further out of reach to teachers, and "link[s] job security [for superintendents] to student achievement and teacher evaluations."[38] A similar proposal recently passed in Michigan—the amendment to the Teacher Tenure Act sparked controversy because it "barred school districts from using seniority as the determining factor when making layoff decisions."[39] The purpose of the amendment, according to one legislator, was to raise the quality of education and to give younger teachers with promising potential a better chance of avoiding layoffs due solely to their lack of experience.[40]

Finally, as discussed below, a recently passed Florida law that ties teacher evaluations to student standardized test scores generated controversy and a lawsuit from teacher unions and individual teachers. Teachers narrowed in on the legal problems with Florida State Board of Education rulemaking procedures and substance. While the law was invalidated by an administrative law judge, there is much room for appeals and further rulemaking.

III. CASE STUDY: FLORIDA

For a number of reasons, Florida's recently passed law is an appropriate one to study in order to understand how this method of teacher evaluation may affect education systems and create ongoing litigation and conflict. The law enacted a value-added model of teacher evaluations based on student test scores and used its new teacher evaluation model as a way to apply for Race to the Top funds.[41] This

---

[35] *Id.*

[36] *Id.*

[37] Bill Barrow, *Teachers File State Lawsuits Challenging Gov. Jindal's Voucher, Tenure Laws*, TIMES-PICAYUNE, June 7, 2012, http://www.nola.com/politics/index.ssf/2012/06/teachers_file_state_lawsuits_c.html.

[38] *Id.*

[39] Lori Higgins, *Teacher Union Files Lawsuit over Michigan Teacher Tenure Act*, DETROIT FREE PRESS, Mar. 8, 2012, www.freep.com/article/20120308/NEWS05/203080731/Teachers-union-files-lawsuit-over-Michigan-Teacher-Tenure-Act.

[40] *Id.*

[41] Postal, *supra* note 28.

legislation also invited controversy and led to a litigation battle.[42] As this section will detail, the litigation created conflict between teachers and the education reform movement, invalidated a rule, and caused uncertainty about future law and policy.

A. The Florida Statute

In 2011, the Florida legislature passed Senate Bill 736, codified in § 1012.34 of the Florida Statutes.[43] Entitled the "Student Success Act," the law provoked a legal challenge from the teachers' union, the Florida Education Association.[44] The statute changes teacher evaluation requirements.[45] Also, "[t]he bill reinforces Race to the Top, which requires 50[%] of the evaluation for classroom teachers and other instructional personnel to be based on student performance for students assigned to them over a [three]-year period."[46] If less than three years of data is available, this part of the evaluation still must account for not less than 40% of the evaluation.[47] The law states that the "Commissioner of Education would establish a learning growth model . . . to measure the effectiveness of a classroom teacher . . . based on what a student learns."[48] This model would be used for the state assessments.[49] The portion of the teacher evaluation measures that are not based on test scores are based on instructional practice.[50] Four evaluation levels are required under the law: "highly effective; effective; needs improvement or, for instructional personnel in the first [three] years of employment who need improvement, developing; and unsatisfactory."[51] The Commissioner of Education is required to consult with experts, instructional personnel, school administrators, and education stakeholders in developing criteria for these levels.[52]

Teachers evaluated as "highly effective" or "effective" will receive a pay raise under the system.[53] A school district will not renew a teacher's annual contract if the teacher receives any of the following: two unsatisfactory annual ratings in a row, two unsatisfactory ratings in a

---

[42] Peek v. State Bd. of Educ., No. 12-1111RP, 2012 WL 3645094 (Fla. Div. Admin. Hearings Aug. 22, 2012)

[43] FLA. STAT. § 1012.34 (2012); Gina Jordan, *Challenge to Teacher Merit Pay Rule Underway in Tallahassee*, STATEIMPACT (May 30, 2012, 12:36 PM), http://stateimpact.npr.org/florida/2012/05/30/challenge-to-teacher-merit-pay-rule-underway-in-tallahassee/.

[44] Jordan, *supra* note 43.

[45] *Id.*

[46] FLA. SENATE COMM. ON EDUC. PRE-K - 12, *supra* note 26, at 1.

[47] FLA. STAT. § 1012.34(3)(a)1a (2012).

[48] *Id.*

[49] *Id.*

[50] *Id.* § 1012.34(3)(a)2.

[51] *Id.* § 1012.34(2)(e).

[52] *Id.*

[53] *Id.* § 1012.22(1)(c)5b.

three-year period, three needs improvement ratings in a row, or a combination of unsatisfactory and needs improvement evaluations in a three-year period.[54] In addition, these performance evaluation ratings can subject a teacher to dismissal for cause.[55] The law also requires superintendent recommendations for promotions or transfers to be primarily based on the employee's effectiveness according to the standards of § 1012.34.[56]

Many Florida teachers reacted negatively to the new law and its requirements that they would have to follow, believing that some measures would not help further educational goals; some who agreed with its basic tenets thought it was implemented too quickly.[57] The research organization that worked to create the policy also mentioned concerns that some districts implemented the plan too quickly.[58] Teachers unions challenged the most controversial part of the law in court—that which uses student test scores as part of the measure of teacher performance.[59]

Another part of the Florida law that generated controversy is the nature of the formula that helps determine teacher pay.[60] The formula predicts what a student should score on the state standardized test (the Florida Comprehensive Assessment Test ("FCAT")), and compares the actual score with the predicted score to evaluate teacher performance.[61] While factors such as "class size, attendance, and the [students'] previous [FCAT] scores" are included, poverty is not included in the formula, despite a known link between poverty and success in school.[62] Other objections to the formula are its complexity and its reliance on variables that may not produce accurate results.[63] As one teacher commented, it is "only as good as the variables that you're actually looking out for as well as the test that you're using to measure."[64] This implicitly raises concerns about the reliability of test scores in general to measure student success.

---

[54] *Id.* § 1012.33(3)(b).

[55] *Id.* § 1012.33(1)(a).

[56] *Id.* § 1012.22(1)(e).

[57] Leslie Postal, *Teachers: New Evaluation System 'Artificial,' 'Frustrating,' 'Humiliating'*, ORLANDO SENTINEL, July 13, 2012, http://articles.orlandosentinel.com/2012-07-13/features/os-florida-teacher-evaluations-20120713_1_new-teacher-evaluations-evaluation-plan-teacher-merit-pay-law.

[58] *Id.*

[59] *Id.*

[60] Laura Isensee & Sarah Gonzalez, *Why Poverty Is Not Included in the Mathematical Equation for Teacher Merit Pay*, STATEIMPACT (Feb. 17, 2012, 11:40 AM), http://stateimpact.npr.org/florida/2012/02/17/why-poverty-is-not-included-in-the-mathematical-equation-for-teacher-merit-pay/.

[61] *Id.*

[62] *Id.*

[63] Sarah Gonzalez, *Inside the Mathematical Equal for Teacher Merit Pay*, STATEIMPACT (Feb. 16, 2012, 2:02 PM), http://stateimpact.npr.org/florida/2012/02/16/inside-the-mathematical-equation-for-teacher-merit-pay/.

[64] *Id.*

B. Delegation to Agencies for Rulemaking and History of Petition

According to § 1012.34, at least half of a teacher's performance evaluation "must be based upon data and indicators of student learning growth assessed annually by the statewide assessments . . . ."[65] The other portion of the evaluation "must include indicators based upon each of the Florida Educator Accomplished Practices ("FEAP") adopted by the State Board of Education" and would come from classroom observations by administrators.[66] The FEAPs are found in Rule 6A-5.065.[67] The Review and Approval Checklist for Instructional Personnel and School Administrator Evaluation Systems ("Checklist") provides elements that must be part of the evaluation systems used; the Checklist is required by § (2)(a) of the proposed rule, but the contents of the Checklist are not spelled out within it.[68]

The Florida legislature delegated responsibility for devising some of the new evaluation procedures to the State Board of Education ("SBE"), the Department of Education ("DOE"), and the Commissioner of Education.[69] The legislature declared that the SBE must adopt rules that "establish uniform procedures for the submission, review, and approval of district evaluation systems and reporting requirements for the annual evaluation of instructional personnel."[70] The legislature also charged the SBE with ensuring that standards for instructor performance levels are specific and discrete.[71] The SBE must also adopt rules that provide for student learning growth measurement[72] and rules tying student growth standards to instructor evaluation ratings.[73] Finally, the SBE is tasked with creating a "process for monitoring school district implementation of evaluation systems in accordance with this section."[74]

The DOE plays a role by receiving instructors' evaluation results from the school district superintendents,[75] approving each individual school district's instructional personnel evaluation systems,[76] and then monitoring the implementation of evaluation systems.[77] The Commissioner of Education "shall consult with experts, instructional

---

[65] FLA. STAT. § 1012.34(3)(a)1 (2012).

[66] *Id.* § 1012.34(3)(a)1c2.

[67] FLA. ADMIN. CODE ANN. r. 6A-5.065 (2012).

[68] 38 Fla. Admin. Weekly 804 (Feb. 24, 2012) [hereinafter Checklist].

[69] *See generally* FLA. STAT. § 1012.34 (2012).

[70] *Id.* § 1012.34(8).

[71] *Id.*

[72] *Id.* § 1012.34(7)(a). The Board must adopt rules implementing the formula to measure student learning growth that the Commissioner of Education. *Id.*

[73] *Id.* § 1012.34(8).

[74] *Id.*

[75] *Id.* § 1012.34(1)(a).

[76] *Id.* § 1012.34(1)(b).

[77] *Id.*

personnel, school administrators, and education stakeholders in developing the criteria for the performance levels."[78] The Commissioner of Education also must "approve a formula to measure individual student learning growth on the [FCAT]" considering certain factors, including an individual student's prior academic performance.[79]

In *Peek v. State Board of Education*, two individual teachers and the Florida Education Association petitioned for a declaration that proposed Rule 6A-5.030 overreached the power delegated to the DOE by the legislature.[80] In addition, petitioners argued that the Rule was vague, arbitrary, and contained technical jargon that, by law, was too complex.[81] Respondents were the DOE and the SBE.[82] SBE members are appointed by the governor, and the DOE administers and implements education policy under the direction and control of the SBE.[83]

The SBE developed its new evaluation system, proposed Rule 6A-5.030, on September 30, 2011.[84] On February 24, 2012, the respondents gave notice that proposed Rule 6A-5.030 would be adopted at a SBE meeting on March 27, 2012.[85] The petitioners learned that the DOE made changes to the proposed Rule without publishing those changes in the Florida Administrative Weekly (FAW), and then adopted the revised Rule during the March 27 meeting.[86] As of the date of the petition, the SBE had not adopted a rule for a value-added formula or started another rulemaking process to adopt such a formula, despite the newly adopted Rule's requirement to use a value-added formula to find student learning growth, which accounts for half of teacher evaluation scores.[87]

By statute, the formula to measure individual student learning growth must measure a student's growth on the FCAT and should include consideration of individual students' prior academic performance, attendance, disability status, and English language learner status.[88] The law also requires that teachers are "fully informed of the criteria and procedures" of the evaluation before the evaluation occurs.[89]

---

[78] *Id.* § 1012.34(2)(e).

[79] *Id.* § 1012.34(7)(a).

[80] Amended Petition at 1, 6–8, Peek v. State Bd. of Educ., No. 12-1111RP, 2012 WL 3645094 (Fla. Div. Admin. Hearings Aug. 22, 2012).

[81] *Id.*

[82] *Peek*, 2012 WL 3645094, at *2.

[83] *Id.*

[84] Amended Petition, *supra* note 80, at 5.

[85] *Id.*

[86] *Id.* at 5–6.

[87] *Id.* at 6.

[88] FLA. STAT. § 1012.34(7)(a) (2012).

[89] *Id.* § 1012.34(3)(b).

C. Rule and Petition Arguments

The purpose of Rule 6A-5.030 was "to implement changes to Section 1012.34, Florida Statutes, as prescribed in the Student Success Act (SB 736) of 2011," and it was intended "to establish procedures for Department of Education review, approval and monitoring of school district systems of personnel evaluation under Section 1012.34, Florida Statutes."[90] Petitioners alleged that Rule 6A-5.030 contained, in addition to rule text, a thirteen-page checklist that included a two-page glossary of terms for school districts to use when they provide rule compliance evidence to the DOE for approval of districts' personnel evaluation systems.[91] Petitioners also alleged that Rule 6A-5.030 did not include any guidance about "the form, content, quantity, or scope of the documents the DOE [would] accept."[92] In addition, petitioners alleged that the Rule incorporated a document, not included in the proposed rule, called "High Effect Size Indicators (2012)." Furthermore, the Rule did not assign a page or form number to the document and did not specifically or clearly reference the document.[93] Petitioners alleged that the incorporation of those documents created a substantive change to the Rule, and that Section 1 of the Rule did not specify the date that districts must turn in evaluation systems to the DOE for approval.[94] Petitioners further argued that Section 2 of the Rule was not specific about how districts could comply, "allow[ing] total discretion to the DOE regarding what compliance with the APR means."[95]

Petitioners also argued that the Rule required teachers to produce documents about more indicators than those codified in another rule and that the Checklist requirements exceeded the Rule's enabling statute.[96] In addition, they argued that the requirements of the evaluation system included some provisions that were difficult to understand, mandated "feedback processes," and required some teacher proficiency monitoring that was not authorized by law.[97] Petitioners claimed that Section 3 of the Rule lacked guidance about how the DOE would decide whether evaluation systems were complete and acceptable.[98] They also argued that Section 4 of the Rule, while making clear the DOE's potential responses to districts' evaluation systems (approval, conditional approval, denial, and approval rescinded), did not give guidance about

---

[90] Checklist, *supra* note 68, at 804.
[91] Amended Petition, *supra* note 80, at 8.
[92] *Id.*
[93] *Id.*
[94] *Id.* at 8–9.
[95] *Id.* at 9.
[96] *Id.*
[97] *Id.* at 10.
[98] *Id.*

the nature of the documentation that would lead to any of these responses.[99]

Petitioners claimed that Section 5 of the Rule, which addressed later modifications to an approved evaluation system and required DOE approval for a district's modifications, was not authorized by law.[100] They contended that mandates from the DOE in Section 6 of the Rule were not authorized by § 1012.34(1)(a) because the DOE required ten mandatory elements of district evaluation system monitoring, including how teachers would be evaluated and evaluation data to be compiled by districts.[101] Petitioners argued that § 1012.34(1)(a) only required that districts report the resulting category of each teacher's evaluation.[102] They also argued that the mandatory Five Year Continuous Improvement Cycle evaluation system review in Section 6 of the Rule was not authorized by law.[103]

Regarding the Checklist, petitioners contended that the Rule did not give directions for satisfactory documentation regarding the approximately fifty goals or indicators about which each school district must send evidence, leaving too much discretion to those working for the DOE.[104] Petitioners stated that the Checklist also mentioned that compliance guidance would be available on the DOE website, but that guidance could be subject to unexamined changes by the DOE without transparent and consistent rulemaking.[105] Petitioners claimed that the Checklist required districts to train employees about the evaluation system without providing any guidelines.[106] They also argued that school districts must provide compliance documentation that will be subjectively found acceptable or unacceptable by the DOE under the Checklist.[107]

In general, petitioners claimed that the proposed Rule 6A-5.030 violated § 120.52(8) and was "an invalid exercise of delegated legislative authority."[108] Petitioners also claimed that the Rule required school districts to rely on policy regarding the learning growth model and value-added formula provided by the DOE that the Commissioner approved but has not been adopted by the SBE, thus unlawfully requiring districts to rely on policy that had not been adopted as a rule.[109]

In the petitioners' argument about failing to follow rulemaking procedures, they argued that respondents did not follow § 120.54(2)(b)

---

[99] *Id.*
[100] *Id.*
[101] *Id.* at 11.
[102] *Id.*
[103] *Id.*
[104] *Id.*
[105] *Id.* at 11–12.
[106] *Id.* at 12–13.
[107] *Id.* at 13.
[108] *Id.*
[109] *Id.*

when they failed to draft the Checklist and High Effect Size Indicator attachment in readable language, instead using technical jargon and cross-referencing in a confusing way.[110] Respondents also violated the requirements of § 120.54(l)(i) when adding the "High Effect Size Indicators (2012)" document without reference to it in the proposed rule and without page numbers "at the last minute."[111] The petitioners also contended that respondents did not follow § 120.54(l)(i) when they did not incorporate additional materials that the DOE would put on its website by reference into the Rule.[112] Many of the rulemaking procedure arguments were addressed by an administrative decision about the Rule, but the more substantive arguments were not addressed.

D. Administrative Law Decision about Procedural Arguments

The administrative law judge ("ALJ") reviewing the Rule discussed how the Florida Administrative Procedure Act deals with incorporative references in rulemaking.[113] The ALJ explained that incorporative references occur when material outside of the legislation is referenced inside the legislation, and the legislation implies or states that the outside material should be treated as part of it.[114] The Florida Administrative Procedure Act, the ALJ noted, allows incorporative references with some restrictions.[115]

The ALJ found the contested rule to be "an invalid exercise of delegated legislative authority, in its entirety."[116] The ALJ concluded that the Rule was invalid because the respondents materially did not follow the proper rulemaking procedures.[117] The ALJ did not decide on the additional substantive matters brought by the petitioners, because doing so would be an improper advisory opinion.[118]

Even though the ALJ did not decide on some of the issues brought by the petitioners, it would be a useful exercise to examine their other legal arguments in light of Florida law, since they may arise in future rulemaking. The Florida statute that deals with whether an exercise of delegated legislative authority is valid is § 120.52(8) of the Florida Administrative Procedure Act.[119]

---

[110] *Id.*

[111] *Id.*

[112] *Id.* at 14.

[113] Peek v. State Bd. of Educ., No. 12-1111RP, 2012 WL 3645094, at *8 (Fla. Div. Admin. Hearings Aug. 22, 2012).

[114] *Id.*

[115] *Id.*

[116] *Id.* at *24.

[117] *Id.* at *23.

[118] *Id.* at *23–24.

[119] FLA. STAT. § 120.52(8) (2012).

E. Substantive Arguments

The petitioners argued that the Rule exceeded the grant of delegated legislative authority under § 120.52(8)(a).[120] Section 120.52(8)(a) provides that:

> Invalid exercise of delegated legislative authority" means action that goes beyond the powers, functions, and duties delegated by the Legislature. A proposed or existing rule is an invalid exercise of delegated legislative authority if any one of the following applies:
>
> (a) The agency has materially failed to follow the applicable rulemaking procedures or requirements set forth in this chapter;
>
> (b) The agency has exceeded its grant of rulemaking authority, citation to which is required by s. 120.54(3)(a)1.;
>
> (c) The rule enlarges, modifies, or contravenes the specific provisions of law implemented, citation to which is required by s. 120.54(3)(a) 1.;
>
> (d) The rule is vague, fails to establish adequate standards for agency decisions, or vests unbridled discretion in the agency;
>
> (e) The rule is arbitrary or capricious. A rule is arbitrary if it is not supported by logic or the necessary facts; a rule is capricious if it is adopted without thought or reason or is irrational; or
>
> (f) The rule imposes regulatory costs on the regulated person, county, or city which could be reduced by the adoption of less costly alternatives that substantially accomplish the statutory objectives.
>
> A grant of rulemaking authority is necessary but not sufficient to allow an agency to adopt a rule; a specific law to be implemented is also required. An agency may adopt only rules that implement or interpret the specific powers and duties granted by the enabling statute. No agency shall have authority to adopt a rule only because it is reasonably related to the purpose of the enabling legislation and is not arbitrary and capricious or is within the agency's class of powers and duties, nor shall an agency have the authority to implement statutory provisions setting forth general legislative intent or policy. Statutory language granting rulemaking authority or

[120] Amended Petition, *supra* note 80, at 14.

> generally describing the powers and functions of an agency shall be construed to extend no further than implementing or interpreting the specific powers and duties conferred by the enabling statute.[121]

The petitioners argued that the Rule should be struck down based on this statute for three primary reasons: (1) the Rule exceeded its grant of authority and enlarged, modified, or contravened the statute; (2) the Rule is vague and lacks adequate standards; and (3) the Rule is arbitrary and capricious. This Note will now examine each of these arguments in turn.

1. The Rule Exceeded Grant of Rulemaking Authority and Enlarged, Modified, or Contravened the Statute

Because arguments challenging an agency's rulemaking actions under § 120.52(8)(b) and § 120.52(8)(c) often overlap, these two sections of the statute will be discussed together.

Under Florida law, an agency that promulgates rules is delegated power to make rules, but that power is limited by the statute providing for the authority.[122] In a case about establishing high water marks by the Board of Surveyors ("Board"), Florida's First District Court of Appeal found that the only authority granted to the Board by statute was to create rules with technical standards so that high water mark surveys are accurately measured for real property boundaries.[123] The legislative grant of authority did not give the Board the authority to define a fixed high water line that would create new legal consequences.[124] The court held that an executive agency is constrained by the legislative grant of authority, even though it may "properly restate or interpret pertinent case[ ]law."[125]

Amendments to Florida's Administrative Procedure Act in the 1990s clarified the limitations imposed on a rulemaking agency. As one court put it:

> Under the 1996 and 1999 amendments to the [Administrative Procedure Act], it is now clear, agencies have rulemaking authority only where the Legislature has enacted a specific statute, and authorized the agency to implement it, and then only if the (proposed) rule implements or interprets specific powers or duties, as opposed to improvising in an area that

---

[121] FLA. STAT. § 120.52(8) (2012).

[122] Bd. of Trs. of Internal Improvement Trust Fund v. Bd. of Prof'l Land Surveyors, 566 So. 2d 1358, 1360 (Fla. 1st Dist. Ct. App. 1990).

[123] *Id.* at 1361.

[124] *Id.*

[125] *Id.*

> can be said to fall only generally within some class of powers or duties the Legislature has conferred on the agency.[126]

The 1999 amendments came as a response to the court's interpretation of the 1996 amendments in *St. Johns River Water Management District v. Consolidated-Tomoka Land Co.*,[127] which interpreted the 1996 amendments to mean that a proposed rule need only be within the range of statutorily granted powers or in the class of powers and duties named in the statute.[128] After the 1999 amendments, a rule created by an agency must be "within the class of powers and duties delegated to the agency," but that alone is not sufficient for a valid exercise of legislative power.[129] An administrative agency's authority when promulgating rules must come from "an explicit power or duty" granted in the statute by the Florida legislature.[130] While a rule will likely be more detailed and specific than the statute giving the authority for the rule, the issue is "whether the statute contains a specific grant of legislative authority for the rule, not whether the grant of authority is specific *enough*."[131]

The First District Court of Appeal used this reasoning to find that a rule prohibiting the use of submerged lands for gambling vessels was invalid because, despite being framed as authorized by statute to regulate anchoring or mooring, it was truly a rule about a commerce-related transport, which the Board was explicitly not allowed to regulate according to a statute.[132] The Court stated that the test under § 120.52(8)(c) "is whether a (proposed) rule gives effect to a 'specific law to be implemented,' and whether the (proposed) rule implements or interprets 'specific powers and duties.'"[133] Without specific powers or duties allowing regulations, the proposed rule is an invalid exercise of delegated legislative authority.[134]

Because an agency's rulemaking authority is delegated from the legislature, an agency may not "enlarge, modify or contravene the provisions of a statute" when creating an administrative rule; otherwise the rule is an invalid exercise of delegated legislative authority.[135] For example, in a case about the Division of Alcoholic Beverages and Tobacco's authority to license restaurants to sell alcohol, the First District Court of Appeal found that the agency "enlarged upon the

---

[126] State v. Day Cruise Ass'n, Inc., 794 So. 2d 696, 700 (Fla. 1st Dist. Ct. App. 2001) (citation omitted).

[127] 717 So. 2d 72, 80 (Fla. 1st Dist. Ct. App. 1998).

[128] *Day Cruise Ass'n*, 794 So. 2d. at 699–701.

[129] Sw. Fla. Water Mgmt. Dist. v. Save the Manatee Club, 773 So. 2d 594, 599 (Fla. 1st Dist. Ct. App. 2000) (quoting FLA. STAT. § 120.52(8) (1999)).

[130] *Id.*

[131] *Id.*

[132] *Day Cruise Ass'n*, 794 So. 2d at 702.

[133] *Id.* at 704 (quoting FLA. STAT. § 120.52(8) (1999)).

[134] *Id.*

[135] State v. Salvation Ltd., 452 So. 2d 65, 66 (Fla. 1st Dist. Ct. App. 1984).

statutory criteria."[136] The legislature had provided for specific criteria to obtain a special restaurant beverage license, including requirements that the business be a restaurant, requirements about size of service area and the number of people that could be served, and percentage of revenue from non-alcoholic products.[137] However, the agency enlarged these requirements by adding that food must be prepared and cooked on the premises.[138]

Another example of a court overturning a regulation because a rulemaking body overstepped its authority is *Department of Health and Rehabilitative Services v. Florida Psychiatric Society*.[139] In this case, the court had to decide whether the Department of Health and Rehabilitative Services had the authority from the legislature to create rules to establish, regulate, and license "crisis stabilization" and "intensive residential treatment" care centers under the Florida Mental Health Act.[140] The court held that the statute that gives the department the authority to protect public health in Florida was explicit in the department's lack of authority to require a license, absent any license requirement specifically provided for by law.[141] In addition to lack of authorization, the court found that the rule created by the department attempted to amend the statute "by diverting persons eligible for services and treatment under the Act into an entirely different evaluation and treatment program."[142] In invalidating the rule, the First District Court of Appeal reiterated that "[a]dministrative regulations must be consistent with the statutes under which they are promulgated, and they may not amend, add to, or repeal the statute."[143]

In a case involving the SBE, however, the court found that a rule promulgated by the agency regulating the revocation of teachers' certificates for community college instruction was valid.[144] The enabling statute provided that instructor employment at community colleges would occur "upon recommendation of the president, subject to rejection for cause by the board of trustees and Subject [sic] to the rules and regulations of the state board relative to certification, tenure . . . and such other conditions of employment as the Division of Community Colleges deems necessary and proper . . . ."[145] The hearing officer found that the SBE only had power to issue certificates, not revoke them, based on the lack of express language about revocation.[146] The court disagreed with

---

[136] *Id.*
[137] *Id.*
[138] *Id.*
[139] 382 So. 2d 1280 (Fla. 1st Dist. Ct. App. 1980).
[140] *Id.* at 1282.
[141] *Id.* at 1283.
[142] *Id.* at 1284.
[143] *Id.* at 1285.
[144] State Bd. of Educ. v. Nelson, 372 So. 2d 114, 115–16 (Fla. 1st Dist. Ct. App. 1979).
[145] *Id.* at 114 (quoting a 1979 Florida statute now codified at FLA. STAT. § 240.335 (2012)).
[146] *Id.*

the hearing officer, instead finding that the SBE did not exceed its authority because its "power to issue certificates in this instance carries with it the power to specify the terms and conditions of their issuance, as well as the terms and conditions upon which the same may be held by the employee, or revoked by the board."[147]

By contrast, the same court found that an agency rule improperly expanded a statute in a chiropractic licensing case.[148] The Department of Professional Regulation (Board of Chiropractic) amended a rule to require a showing of professional accreditation, in addition to regional accreditation, for the chiropractic college from which a licensing applicant graduated.[149] The court found, however, that the statute only required that an applicant for licensure be a graduate of a chiropractic college accredited by an agency approved by the U.S. Department of Education and the Council on Postsecondary Accreditation.[150] The rule's additional requirements of regional and professional accreditation improperly expanded the statute.[151]

An agency also may not contravene the provisions of a statute when creating a rule.[152] For example, the First District Court of Appeal held that the Agency for Persons with Disabilities contravened the plain language of the statute by creating age requirements for certain Medicaid waiver benefits.[153] The statute at issue said only that clients eligible for those benefits "shall include, but is not limited to, clients requiring residential placements, clients in independent or supported living situations, and clients who live in their family home."[154] The court held that by adding age restrictions, the agency contravened the statute, despite the ALJ's finding that the rule was justified by the agency and was therefore valid.[155] The court emphasized that despite the agency's policy justification of its rule, a rule that contravenes a statute is nonetheless invalid.[156] The agency also contravened the statute by creating a rule that would assign a client's level of benefits based only on whether they previously received benefits, when the statute required assignment based on an individual assessment.[157] The court held that this requirement contravened the express language in the statute by replacing requirements, whereas the age requirement contravened the statute by

---

[147] *Id.* at 114–15.

[148] Dep't of Prof'l Regulation v. Sherman Coll. of Straight Chiropractic, 682 So. 2d 559 (Fla. 1st Dist. Ct. App. 1995).

[149] *Id.* at 559–60.

[150] *Id.* at 561.

[151] *Id.*

[152] FLA. STAT. § 120.52(8)(c) (2012).

[153] Moreland *ex rel.* Moreland v. Agency for Pers. with Disabilities, 19 So. 3d 1009, 1010–12 (Fla. 1st Dist. Ct. App. 2009) (quoting FLA. STAT. 393.0661(3)(c) (2007)).

[154] *Id.*

[155] *Id.* at 1012.

[156] *Id.*

[157] *Id.*

adding a factor to the statute about whether an individual could receive a certain level of benefits.[158]

Even when it is possible that the legislature mistakenly neglected to give authority to an agency, as evidenced by overall legislative policy, or when an administrative rule attempts to reach an admirable goal beyond its authority, the courts cannot rewrite statutes by allowing agencies to enlarge or contravene statutory authority.[159] Similarly, rulemaking authority cannot be accorded to an agency, even though the public might benefit, if the legislature has not given the agency that authority.[160]

In *Peek*, petitioners argued that some parts of the Rule proposed by the DOE went beyond what the Florida legislature authorized the agency to do, thus exceeding its grant of authority in violation of § 120.52(8)(b) and enlarging, modifying or contravening the agency's enabling statute in violation of § 120.52(8)(c).[161] For example, petitioners argued that documentation requirements regarding evaluation systems went beyond the FEAP indicators required by statute.[162] Part of the Rule provided that evaluation systems must "[c]ontain evidence of each of the elements as described in the Review and Approval Checklist for Instructional Personnel and School Administrator Evaluation Systems, Form No. EQEVAL-2012."[163]

Other examples of requirements where petitioners claimed that the agency went beyond what the legislature authorized were the Five Year Continuation Improvement Cycle and the DOE evaluation system quality control monitoring process.[164] Petitioners alleged that the Five Year Cycle would require DOE review of district evaluation systems every five years. while the enabling statute only required that district school boards monitor their own evaluation systems annually and authorized DOE assistance only at the district's request.[165] In addition, petitioners argued that the DOE had taken too much authority in the Rule by requiring evaluation system modifications to go through the agency for approval before implementation.[166]

According to the court's interpretation, the DOE can create rules that are more specific than the enabling statute, but it should not create rules with new legal consequences that are not part of an explicit duty given to them through the statute. The rule's validity requires specific grants of authority, not merely the DOE acting within a general class of powers. Because this was made clear from the 1999 amendments that

---

[158] *Id.*

[159] Capeletti Bros., Inc. v. Dep't of Transp., 499 So. 2d 855, 857 (Fla. 1st Dist. Ct. App. 1986).

[160] Witmer v. Dep't of Bus. & Prof'l Regulation, 662 So. 2d 1299, 1302 (Fla. 1st Dist. Ct. App. 1995).

[161] Amended Petition, *supra* note 80, at 14–15.

[162] *Id.* at 16.

[163] Checklist, *supra* note 68, at 805.

[164] Amended Petition, *supra* note 80, at 15.

[165] *Id.* at 16–17.

[166] *Id.* at 10.

reacted to the court's interpretation of previous amendments, the petitioners in *Peek* have a strong argument that parts of the DOE's Rule are invalid. The statute does not mention what would happen if districts modified their evaluation systems, but the Rule provides for submittal to the DOE when a system has been modified substantially. Adding a section of a rule not provided for by the statute is arguably invalid under both § 120.52(8)(b) and under *State v. Day Cruise Association.*[167] The *Day Cruise Association* court reasoned in a straightforward example that if an agency's rule utilizes a power specifically denied to that agency, then the rule is invalid.[168]

Even in less straightforward examples, Florida courts have found that an agency cannot expand its powers. Specifically, in *State v. Salvation Limited*, the court found enlarged statutory criteria when an agency added requirements that food must be prepared on the premises of restaurants.[169] While there was nothing that prohibited this requirement in the statute, the court held that the agency could not expand its power to require restaurants to comply with a rule unauthorized by statute.[170] A similar enlargement of specific provisions of law arguably occurred in § 3(a)(2) of the teacher evaluation statute, which required use of FEAPs as indicators for teacher observations. The Rule added the Checklist with additional indicators; this addition could arguably be found by a court as enlarging an agency's authority beyond the statute.

Florida courts have also looked to the effects of regulations. For example, a court held in *Department of Health v. Florida Psychiatric Society*, not only that creation of new care centers by the agency had been done without statutory authority, but that the regulations, in effect, amended the statute by diverting clients from authorized treatment centers into different treatment programs.[171] Similarly, in *Department of Professional Regulation v. Sherman College of Straight Chiropractic*, the agency was not authorized to add regional and professional accreditation requirements when the statute only called for accreditation generally.[172] The issues in these cases look similar to those posed by the Rule's Five Year Continuous Improvement Cycle; while review of evaluation systems was provided for in the statute, a five-year cycle was not included. The inclusion of the five-year cycle provision arguably has the effect of rewriting the enabling statute, which would be prohibited under *Florida Psychiatric Society* and *Sherman College.*

One example of a case in which an agency rule was upheld is *State Board of Education v. Nelson*, which allowed the SBE the power to

---

[167] 794 So. 2d 696 (Fla. 1st Dist. Ct. App. 2001).
[168] *Id.* at 697.
[169] 452 So. 2d 65, 66 (Fla. 1st Dist. Ct. App. 1984).
[170] *Id.* at 66–67.
[171] 382 So. 2d 1280, 1285 (Fla. 1st Dist. Ct. App. 1980).
[172] 682 So. 2d 559 (Fla. 1st Dist. Ct. App. 1995).

revoke community college certificates as a part of the statute allowing the agency to issue certificates.[173] While this case also involved the DOE—the same agency that promulgated the Rule—*Nelson* was decided before the 1999 amendments, which restricted the rules agencies could create to those pursuant to specific delegated powers. Because the *Nelson* holding was premised on a since-amended state statute, it is unlikely that a court would draw on *Nelson* to uphold the Rule.

In summary, Florida case law supports many of the petitioners' arguments about the invalidity of the Rule pursuant to § 120.52(8)(b) and § 120.52(8)(c). While it is not known how a court would rule on these arguments, a court considering them in subsequent litigation may find them compelling. If new rules are promulgated that include the parts of the Rule that were not explicitly invalidated by the ALJ for procedural reasons in the earlier proceeding, it is possible that the above substantive arguments might render a new rule invalid as well.

2. Vague and Lacks Adequate Standards

The Florida Administrative Procedure Act defines invalid exercises of legislative authority to include rules that are vague, fail to establish adequate standards for agency decision-making, or vest unbridled discretion in the agency.[174] Florida courts have held that a regulation is vague "if it 'either forbids or requires the doing of an act in terms so vague that men of common intelligence must necessarily guess at its meaning and differ as to its application.'"[175] When terms are used in a rule that is not defined, the "common ordinary meaning" applies.[176] In *Cole Vision Corporation*, the First District Court of Appeal addressed whether the terms "'implies or suggests,' 'clearly and sufficiently indicate,' and 'associated or affiliated with'" were so vague as to render the regulation in which they were used invalid.[177] The court held that these terms were easily understood by an ordinary person, especially if the terms were read in conjunction with the other parts of the rule.[178]

In addition to challenging the regulations on vagueness grounds, the *Cole Vision Corporation* appellants questioned whether the rule created "adequate standards to assist optometrists in determining whether they [we]re engaged in a business relationship with a commercial

---

[173] 372 So. 2d 114 (Fla. 1st Dist. Ct. App. 1979).

[174] FLA. STAT. § 120.52(8)(d) (2012).

[175] Cole Vision Corp. v. Dep't of Bus. & Prof'l Regulation, 688 So. 2d 404, 410 (Fla. 1st Dist. Ct. App. 1997) (quoting Witmer v. Dep't of Bus. & Prof'l Regulation, 662 So. 2d 1299, 1300 (Fla. 1st Dist. Ct. App. 1995)).

[176] *Id.*

[177] *Id.*

[178] *Id.*

establishment in violation of [the statute]."[179] The court also rejected this challenge on the basis that "excruciating detail" in the rule was not necessary to determine all contingencies about how the business relationship might occur in optometry, and that the rule established adequate standards.[180]

Florida case law contains other examples of rules challenged on grounds of vagueness, inadequate standards, or unfettered discretion. In *Witmer v. Department of Business and Professional Regulation*, the Fourth District Court of Appeal held that an agency's rule prohibiting corrupt or fraudulent practices in horse and dog racing was so vague that it was invalid.[181] The rule, the court noted, "punishe[d] corrupt or fraudulent practices without ever defining them or referring to a standard by which a practice may be judged to be corrupt or fraudulent."[182] By contrast, in *Humhosco v. Department of Health and Rehabilitative Services*, the court found that the terms "normally" and "substantially" were terms that are "commonly used and understood" and were within the agency's discretion to interpret and use on a case-by-case basis, thereby upholding the regulations against a challenge that the regulations violated the state constitution by giving the agency unbridled discretion.[183]

In *Cortes v. State Board of Regents*, the challenged rule gave new "unguided discretion" to agencies.[184] The court held that if rules create discretion that is not contained in the statute giving rulemaking authority, the rule must provide a basis for the discretion.[185] Without a basis for the discretion, the rule lacks sufficient standards, and is thus deemed invalid.[186]

Petitioners claimed that the Rule violated § 120.52(8)(d) because it was vague, lacked adequate standards, and gave too much discretion to the DOE.[187] The Rule required that evaluation systems submitted to the DOE demonstrate promotion of "continuous improvement of student learning growth and faculty and leadership development through feedback processes."[188] It stated several procedures that must be included, such as "procedures to ensure rater accuracy and reliability, training of employees on proficiency expectations, and monitoring of

---

[179] *Id.*

[180] *Id.*

[181] Witmer v. Dep't of Bus. and Prof'l Regulation, 662 So. 2d 1299, 1300 (Fla. 1st Dist. Ct. App. 1995).

[182] *Id.* at 1302.

[183] Humhosco, Inc. v. Dep't of Health & Rehab. Servs., 476 So. 2d 258, 261 (Fla. 1st Dist. Ct. App. 1985).

[184] Cortes v. State Bd. of Regents, 655 So. 2d 132, 138 (Fla. 1st Dist. Ct. App. 1995).

[185] *Id.*

[186] *Id.*

[187] Amended Petition, *supra* note 80, at 17.

[188] Instructional Personnel and School Administrator Evaluation Systems, *supra* note 163, at 805.

improvement results in student learning growth and instructional personnel and school leader proficiency on evaluation indicators."[189]

Petitioners argued that the Rule gave the DOE too much discretion to approve or reject evaluation systems that have high-stakes evaluation outcomes for teachers in Florida.[190] Petitioners also argued that the fifty indicators that help determine whether evaluation systems comply with the agency's rules were not accompanied by information about what quantity or quality of documentation is required for approval or how a system that is not approved can be remedied.[191] In addition, petitioners alleged that the DOE intended to modify the Rule with materials on its website, but failed to incorporate them by reference into the Rule and failed to ensure that the website could not be changed at the discretion of the DOE over time.[192] These issues, petitioners alleged, showed that the Rule was vague, lacked adequate standards, and gave unbridled discretion to the DOE.[193]

A court would likely be less receptive to these arguments about vagueness and lack of standards as compared to the enlarging, modifying, and contravening argument above. In *Peek*, petitioners argued that the terms "procedures to ensure," "training employees," and "monitoring of improvement results" were too vague.[194] But Florida courts seem to require an utter lack of standards to find that terminology is too vague for a common person to understand. Terms such as "normally," "substantially," "clearly and sufficiently indicate" have been approved by the courts, and complete details are not required. In turn, courts are less likely to invalidate a rule on vagueness grounds and are more likely to be persuaded that a rule lacks ascertainable standards. Petitioners' argument that the Rule does not specify which indicators are required to be documented may be more persuasive to a Florida court, because districts will not know how much or what kind of information they need to provide to comply.

In summary, while petitioners' arguments about the invalidity of the Rule on vagueness and inadequate standards grounds are not as strong as their arguments that the Rule is outside the legislative grant of authority (as discussed in the previous section), they are still potential arguments that could lead to further litigation if a new rule were promulgated in the future.

---

[189] *Id.*

[190] Amended Petition, *supra* note 80, at 17.

[191] *Id.* at 18.

[192] *Id.*

[193] *Id.* at 17–18.

[194] *Id.* at 9–10.

3. Arbitrary and Capricious

An agency cannot promulgate a rule that is arbitrary or capricious, under § 120.52(8)(e).[195] Under that provision, "a rule is arbitrary if it is not supported by logic or the necessary facts."[196] It "is capricious if it is adopted without thought or reason or is irrational."[197]

Florida courts' analysis of whether a rule is arbitrary and capricious is similar to rational-basis review of legislation under the Equal Protection Clause of the U.S. Constitution.[198] An agency may interpret the authority-granting statute within a range of possible interpretations and does not have to pick the most desirable interpretation.[199] The court in *Florida League of Cities* stated that "if an enabling statute . . . simply states that an agency may 'make such rules and regulations as may be necessary to carry out the provision of this act,' the regulations promulgated thereunder will be sustained as long as they are reasonably related to the purposes of the enabling legislation, and are not arbitrary or capricious."[200] Under this standard, rules must not be irrational or illogical.[201] The court reasoned that the rule in question, which would make "a wastewater treatment facility liable for proper disposal of its domestic wastewater residuals" with a few stated exceptions, was authorized by statutes that gave the Department of Environmental Regulation the power to create rules regarding control and prohibition of air and water pollution.[202] The court found that the rule's requirements appropriately helped reach the goals as put into the statute since residuals pollute air and water.[203]

F. Summary of Substantive Argument Analysis

Petitioners in the teacher evaluation case claimed that the Rule is arbitrary and capricious, and thus an invalid exercise of delegated legislative authority under § 120.52(8)(c), because it does not include "defined standards for initial evaluation system approval" or standards for its Five Year Continuous Improvement Cycle Monitoring.[204]

---

[195] FLA. STAT. § 120.52(8)(e) (2012).

[196] *Id.*

[197] *Id.*

[198] Fla. League of Cities, Inc. v. Dep't of Envtl. Regulation, 603 So. 2d 1363, 1369 (Fla. 1st Dist. Ct. App. 1992).

[199] *Id.*

[200] *Id.* (citing Fla. Waterworks Ass'n v. Fla. Pub. Serv. Comm'n, 473 So. 2d 237, 239–40 (Fla. 1st Dist. Ct. App. 1985)).

[201] *Id.* at 1367.

[202] *Id.* at 1370.

[203] *Id.*

[204] Amended Petition, *supra* note 80, at 18–19.

Petitioners also claimed that the Rule created a very broad evaluation system review by the DOE and created a process for monitoring that lacks logic and reason.[205]

Even though petitioners argued that the Rule about teacher evaluations is arbitrary and capricious, since the Rule is reasonably related to the goal of the statute, a court may not be convinced by this argument. Rational basis review does not require the best policy to be chosen, and the arbitrary and capricious standard is a low one for an agency to meet. This argument is the weakest of the petitioners' three legal arguments discussed in this Note, but since the ALJ did not rule on it previously, it could be raised again if a similar rule were promulgated.

IV. PREVENTING LITIGATION

Although the Rule was invalidated on procedural rather than substantive grounds, the ALJ pointed out that subsequent rulemaking may occur.[206] While the ALJ did not rule on the substantive issues in the petition, he mentioned that these arguments are not foreclosed if further rulemaking is attempted.[207] Further rulemaking could lead to another challenge by teachers unions, unless they are brought into the discussion to find a common solution. Since at least some of the substantive arguments raised by petitioners (especially those claiming that the DOE and SBE exceeded their rulemaking authority under the statute) would likely prevail, the rulemaking process could lead to further litigation and ultimately prolong the process for a long period of time, if the agencies continue to promulgate rules without teacher input or support.

Other states have included teachers in the discussion about teacher evaluation, even sometimes after strife between policymakers and teachers unions. For example, Massachusetts recently passed a law to change teacher evaluation processes. The Massachusetts Teachers Association, Massachusetts's largest teachers union, and Stand for Children, an education advocacy organization, compromised on the issue of teacher seniority.[208] The Massachusetts Teachers Association gave up some seniority rights while Stand for Children opted not to put forward a broader ballot initiative, which included "a more expansive plan to change the way teachers are hired, transferred, and laid off."[209]

---

[205] *Id.* at 19.

[206] Peek v. State Bd. of Educ., No. 12-1111RP, 2012 WL 3645094, at *23–24 (Fla. Div. Admin. Hearings Aug. 22, 2012).

[207] *Id.*

[208] Frank Phillips, *Massachusetts Teachers Union Agrees to Give up Key Rights on Seniority*, BOSTON GLOBE, June 8, 2012, http://www.bostonglobe.com/metro/2012/06/08/massachusetts-teachers-union-agrees-give-many-seniority-rights/GB6B5YhIcriROeDLtULLRI/story.html.

[209] *Id.*

Legislators were brought in and agreed to pass the compromise.[210] The ballot initiative would have based teacher retention and promotion entirely on test scores and performance reviews.[211]

Two of the three stated purposes of the Massachusetts bill were: "to assure the effective implementation of the education evaluation system adopted by the board of elementary and secondary education by providing training for teachers and administrators in evaluation and supervision" and "to assure that indicators of job performance as evidenced by evaluation and other factors are the primary factors in school staffing decisions."[212] The third purpose is about data collection for the evaluation system.[213] The bill provides for evaluation training and funds for training to make the evaluation process effective.[214]

In the bill's provision on how teachers will be evaluated, factors are listed that explain how teachers will be evaluated in various situations.[215] Section 3 of the bill provided that teachers with "professional teacher status" could not be laid off due to "a reduction in force or reorganization if there is a teacher without such status for whose position the covered employee is currently certified or if there is a less qualified teacher with such status holding the same or similar position for which the covered employee is currently certified."[216] While this provision benefitted the teachers with seniority, other parts of the bill provided that collective bargaining agreements could not displace professional status teachers with more senior teachers "unless the more senior teacher is currently certified pursuant to section 38G and is at least as qualified for the position as the junior teacher holding the position."[217] The bill states that the criteria for whether a teacher is qualified must include as primary factors: "indicators of job performance, including overall ratings resulting from comprehensive evaluations conducted consistent with section 38 and the best interests of the students in the school or district."[218] Collective bargaining can only be used in negotiations for seniority or length of service when the teachers involved have the same level of qualifications as specified in the statute.[219]

Another example begins with the Chicago teacher strike at the beginning of the 2012–2013 school year, which included seven days of missed classes, but ultimately ended in a compromise over the issue of

---

[210] *Id.*

[211] Mass. Initaitive Petition, *An Act Prompting Excellence in Public Schools*, No. 11-20 (Aug. 3, 2011), *available at* http://www.mass.gov/ago/docs/government/2011-petitions/11-20.pdf.

[212] S.B. 2315, 187th Gen. Ct. (Mass. 2012), *available at* http://www.malegislature.gov/Bills/187/Senate/S02315/History.

[213] *Id.*

[214] *Id.* at 5.

[215] *Id.* at 3

[216] *Id.*

[217] *Id.*

[218] *Id.*

[219] *Id.*

the teacher evaluation system.[220] Teacher evaluations will be partly based on student test scores, but not to the degree originally proposed.[221] After the strike ended, Randi Weingarten, president of the American Federation of Teachers, stressed the need for public education reform to include teacher and parent input, in order to attain the common goal of improved education for students.[222] The Chicago teacher strike showed a negative consequence similar to those that litigation may bring to education. More significant attempts at compromise and greater inclusion of teachers and education experts before such an episode erupts may be able to decrease the probability that students will suddenly miss several days of school. Chicago teachers were only involved in compromise discussions after the strike, demonstrating that it may be preferable to bring teachers into the conversation earlier in the conflict.

The National Governors Association issued recommendations for how states can implement merit pay policies.[223] According to the recommendations, teacher evaluation systems should be fair and based partly on teacher input, but tests should not be the sole basis of teacher evaluations.[224] The recommendations especially stressed the importance of including teachers in the policymaking process.[225]

In addition to the concerns about increased litigation in response to new merit pay measures in the states, another concern is that pay, evaluations, and hiring decisions based on student test scores may not create incentives for teachers to improve their performance. The purpose of Florida's § 1012.34, which created a teacher evaluation system based on student test scores, is to "increas[e] student learning growth by improving the quality of instructional, administrative, and supervisory services in the public schools of the state."[226] But as discussed below, it is unclear whether having teacher evaluations based on student scores would actually improve quality of instruction.

---

[220] John Byrne & Hal Dardick, *Emanuel Lauds End of Strike*, CHI. TRIBUNE, Sept. 19, 2012, http://articles.chicagotribune.com/2012-09-19/news/ct-met-teachers-strike-rahm-emanuel-0919-20120919_1_new-teacher-evaluation-system-laid-off-teachers-rate-teachers.

[221] *Id.*

[222] Lyndsey Layton, *After Chicago Success, Teachers Unions Spread Their Message*, WASH. POST, Sept. 21, 2012, www.washingtonpost.com/local/education/after-chicago-success-union-leaders-spread-their-message/2012/09/21/cad16290-0424-11e2-9b24-ff730c7f6312_story.html?hpid=z2.

[223] Jason Koebler, *Governors Association Examines Teacher Merit Pay*, U.S. NEWS, Jan. 9, 2012, www.usnews.com/education/blogs/high-school-notes/2012/01/09/governors-association-examines-teacher-merit-pay.

[224] *Id.*

[225] *Id.*

[226] FLA. STAT. § 1012.34(1)(a) (2012).

V. PUBLIC POLICY

A. Efficacy of Tying Teacher Evaluations or Pay to Student Scores

Various studies and policy articles have identified several problems with using student test scores to measure teacher performance, even with a value-added model. First, value-added models have measurement problems to work out, and until researchers do more work, this method of accountability should not be used on its own. Second, standardized tests emphasize low-level skills rather than critical thinking; relying on tests to evaluate teachers will not highlight those educators who teach critical thinking skills well. Third, teachers who have been scored based on their students' testing performance often receive fluctuating evaluation results from year to year, calling into question the tying of teacher performance to student standardized test scores. Finally, student achievement did not appear to increase after a teacher merit pay system was introduced in the New York area, and similar results appeared in a Vanderbilt study about merit pay, questioning the notion that high-stakes consequences for teachers based on student test scores will increase student achievement.[227]

Value-added models of test-based accountability programs still need further development; at this point, the information they provide is "error-prone and has a number of other important limitations."[228] The value-added model is sometimes used to represent a student's "total growth" over the school year rather than how much the teacher's efforts contributed to the student's growth over the year, a distinction that should make a difference when using student scores for teacher accountability.[229] Furthermore, the value-added model may not show student growth when a higher curriculum is used at a particular school. For example, in a seventh-grade math class where one class focuses mostly on algebraic concepts and another focuses on arithmetic, the state test may show more growth from the students focusing on arithmetic if the test does not focus on algebra.[230] Additionally, if student test scores will be used to make employment decisions about individual teachers, sampling error is a concern because of the small sample size of a teacher's limited number of students.[231] In addition, "[t]here has been very little research on the practical effects of using [value-added models.]"[232]

---

[227] *Infra*, notes 240–43.
[228] Koretz, *supra* note 8, at 18.
[229] *Id.* at 19.
[230] *Id.* at 21.
[231] *Id.* at 26.
[232] *Id.* at 39.

Another problem with using student test scores as a way to measure teacher quality is that many standardized tests measure mostly low-level skills, instead of a student's critical analysis skills.[233] Basing teacher evaluations on student test scores may not recognize teachers who excel at teaching these higher-level analysis skills.[234] As such, even though some argue that teacher evaluation systems should be reformed, "there are also good reasons to be concerned about claims that measuring teachers' effectiveness largely by student test scores will lead to improved student achievement."[235]

There are several studies that question the ability of student test scores, even in a value-added model, to adequately measure teacher effectiveness.[236] In a study from the Economic Policy Institute, researchers cautioned against using student test scores as a large percentage of teacher evaluations, citing plans that give 50% evaluation weight to student test scores as too much.[237] The study also noted that "there is broad agreement among statisticians, psychometricians, and economists that student test scores alone are not sufficiently reliable and valid indicators of teacher effectiveness to be used in high-stakes personnel decisions."[238] Especially worrisome is the fluctuation among teachers' ratings from year to year; for example, in one study, one-third of teachers ranked in the top 20% of effectiveness in one year moved to the bottom 40% in effectiveness the following year.[239] Research organizations such as the Board on Testing and Assessment of the National Research Council of the National Academy of Sciences, the Educational Testing Service's Policy Information Center, and the RAND Corporation have cautioned that even value-added models of evaluation are unstable, imprecise, and should not be considered fair or reliable.[240]

In a study conducted by Vanderbilt University in Nashville public schools, offering teachers up to $15,000 in return for improved student test scores did not change academic performance in a discernible way.[241] The professor who led the study advocated developing "more thoughtful and comprehensive ways of thinking about compensation."[242] In a study conducted by Harvard economist Roland Fryer, it was found that New York City's merit pay system did not increase student achievement, and

---

[233] Toch, *supra* note 2, at 35.

[234] *Id.*

[235] Baker, *supra* note 3, at 1.

[236] *Id.* at 2.

[237] *Id.*

[238] *Id.*

[239] *Id.*

[240] *Id.* at 2–3.

[241] Nick Anderson, *Teacher Bonuses Not Linked to Better Student Performance, Study Finds*, WASH. POST, Sept. 21, 2010, www.washingtonpost.com/wpdyn/content/article/2010/09/21/AR2010092103413.html.

[242] *Id.*

achievement seemed to decline at the middle school level.[243] The experiment in merit pay targeted high-need schools and gave teacher bonuses based on the schools' performance as measured by the city's progress report cards.[244]

For policy efficacy reasons, merit pay and value-added models should be more carefully examined before tying much federal education funding to these models. As this Note has shown, education experts (not only teachers unions) are calling for a closer look, if not a complete reversal, of these policies.

B. Work Environment for Teachers

Another potential negative consequence of relying on student test scores for teacher evaluations is that teaching to the test policies may lead the most effective teachers to seek employment in another field. A study found that testing in schools takes away from opportunities for teachers to be creative, and that was what most new elementary school teachers felt was a major drawback of teaching.[245] It was the third most popular reason for new secondary school teachers leaving, behind discipline or behavioral issues and lack of student motivation.[246] High-stakes testing tied to teacher evaluations might create another reason for education systems to focus on teaching to the test, and might negatively affect teacher autonomy and creativity, which could potentially drive away teachers who enjoy teaching higher-level analysis skills from the profession.

VI. RECOMMENDATIONS

Florida's experience in enacting a statute and regulation that require at least 50% of teacher evaluations to be based on student test scores can help inform other states thinking of implementing these policies. However, since at least a 50% reliance on student test scores is required to receive Race to the Top funds, a change in national policy may be required, since state policymakers may feel that they have no choice but to compete for Race to the Top funds. The Florida experience involved

---

[243] Elizabeth Green, *Study: $75M Teacher Pay Initiative Did Not Improve Achievement*, GOTHAM SCHOOLS, Mar. 7, 2011, http://gothamschools.org/2011/03/07/study-75m-teacher-pay-initiative-did-not-improve-achievement/.

[244] *Id.*

[245] JONATHAN ROCHKIND ET AL., NAT'L COMPREHENSIVE CTR. FOR TEACHER QUALITY AND PUB. AGENDA, LESSONS LEARNED: NEW TEACHERS TALK ABOUT THEIR JOBS, CHALLENGES, AND LONG-RANGE PLANS 13 (2007), *available at* http://www.publicagenda.org/files/pdf/lessons_learned_1.pdf.

[246] *Id.*

much conflict between the parties in public education and education policymaking. This conflict led to litigation invalidating the Rule necessary to implement the new law. Not only will the invalidated Rule likely necessitate new rulemaking, the administrative law judge who invalidated the Rule left many substantive legal questions unresolved. Because these questions remain unresolved, it is possible that a new rule will be invalidated based on some of these legal issues, and it is also possible that decisions will be appealed, creating uncertainty for months and years to come.

Florida's experience and difficulty with creating a new teacher evaluation system may counsel other states to approach policymaking by including teachers in the lawmaking process. Massachusetts is an example where teachers and policymakers who disagreed on policy came together in a compromise, allowing both sides to make a step toward improving the education system in a way that comported with their viewpoints. Such compromise may reduce the amount that states rely on student test scores in evaluating teachers.

While concerns about ongoing conflict and litigation were key to the problems that occurred in Florida, student test scores are likely not the best solution for formal teacher evaluations, even when using value-added models. Using student test scores may help principals and teachers improve their schools by providing information on student performance in certain subjects, but these scores only present one aspect of the educational process and only test certain skills. As discussed above, value-added models examined further before using them in a high-stakes way; determining which variables to include in the value-added model and taking time to examine sampling error and differences in curricula that may affect the model should take place before requiring the models for teacher evaluations. Many education researchers raised, and continue to raise, questions about placing too much value on testing in education and whether emphasizing testing motivates teachers to perform better at work. With such conflicting evidence and studies about the efficacy of using student test scores (even value-added models) to formally evaluate teachers, at the very least, more research should be conducted and considered before enacting these policies. In the education policy community, there simply has not been enough discussion about the strengths and weaknesses of value-added models to begin implementing them in teacher evaluations.[247]

The recommendations from this Note are twofold. Policymakers at the national and state levels should be involved, since much of the recent reliance on student test scores for teacher evaluation has occurred based on Race to the Top. First, policymakers should consider the research about the efficacy of value-added models, since many studies have found that they still have many problems to resolve before using them in a

[247] Koretz, *supra* note 8, at 18–19.

high-stakes way. Second, whether policymakers decide to use value-added models of student test scores or not, they should attempt to include teachers' groups in a meaningful way when creating new systems of teacher evaluation. Litigation and conflict have left Florida uncertain about what rules might be used to implement the evaluation system, and teacher opposition to a method for improving student performance negatively affects implementation of a new system. To reduce this conflict, a genuine inclusion of teacher views, coupled with further research on proposed evaluation systems, will be necessary to improve public education overall. As the case study in Florida has shown, a reliance on student test scores as the primary method of evaluating teachers is neither workable nor beneficial. Other options must be developed.