

COMMISSIONER, NEW JERSEY DEPARTMENT OF EDUCATION

IN THE MATTER OF THE PROCEEDING BETWEEN:

Gerri A. Chapman, “Respondent” or “Tenured Teacher”

- and -

State-Operated School District of the City of Camden, “Petitioner” or “District”

Agency Docket No: 248-9/16

AWARD AND OPINION

Hearings: Oct. 31 and Nov. 1, 2016

Posthearing Briefs: Dec. 5, 2016

Date of Award: Dec. 13, 2016

Arbitrator: Perry A. Zirkel

APPEARANCES:

For the Camden School District

Joseph G. Antinori, Esq.
Brown & Connery, LLP
360 Haddon Avenue
Westmont, NJ 08108

For Tenured Teacher Geraldine Chapman

Cosmas P. Diamantis, Esq.
Zeller & Wieliczko, LLP
120 Haddontowne Court
Cherry Hill, NJ 08034

Factual Findings

On or about Jan. 1, 2001, the District hired Geraldine Chapman as an elementary school teacher.¹ During the next decade, she served as in four elementary schools, attaining tenure during the first part of this period.

Starting in the 2011–12 school year, the District assigned Chapman to teach kindergarten at the Forest Hills Elementary School. She continued at Forest Hills thereafter until the pertinent period. More specifically, in 2014–15, she co-taught a grade 2 inclusion class² with special education teacher Laura Colligan,³ and in 2015–16, she taught a regular grade 2 class on her own.

On June 26, 2013, per a consent order between the Commissioner of Education and the District board of education in March 2016, the District changed to state-operated status.⁴ A primary resulting priority was on improving student outcomes.

In the 2013-14 school year, the District initiated its implementation of the Teacher Accountability for the Children of New Jersey (TEACHNJ) Act.⁵ With Commissioner approval, the District selected the Danielson framework for the practice part⁶ of its evaluation rubric. This framework consists of four domains, each with specified components. The domains are planning

¹ As a matter of legal style, this decision uses after the first, full reference to each individual solely his or her surname without titles or other prefixes and without any intent of discourteousness. Moreover, where the exhibits are identified as “Petitioner” (for those admitted into evidence from the District) and “Respondent” (for those admitted into evidence from the Tenured Teacher in this case), “Petitioner/Respondent” for duplicates, and “Joint” for those submitted jointly. Additionally, “Tr.” refers to specified pages in the first (“-I”) or second (“-II”) volume of the hearing transcript. Finally, the footnotes use “*infra*” and “*supra*” for cross references to parts of this decision respectively below or above that footnote.

² The total number of students in the class was approximately 20, including approximately 5 special education students.

³ Although her surname subsequently changed upon her marriage, this arbitration decision uses this prior surname from the case exhibits for consistency.

⁴ See, e.g., Joint exhibit 1.

⁵ For an overview of the provisions of this legislation, which included a pilot implementation date of Jan. 31, 2013, see *infra* notes 44–54 and accompanying text.

⁶ See *infra* note 52 and accompanying text.

and practice, classroom environment, instruction and student engagement, and professional responsibilities. The District provided for training for principals concerning implementation of the Danielson framework, including computerized documentation via the Teachscape digital platform, in December 2012 and spring 2013, with the principals providing training to their teachers in September 2013 and periodic follow-up meetings in fall 2013.

For implementation of the student learning⁷ part of the evaluation rubric in the affected elementary grades in 2014-15,⁸ the District obtained approval for the Northwest Evaluation Association's (NWEA's) Measures of Academic Performance (MAP) assessment for the student learning⁹ part of the evaluation rubric, specifically for what New Jersey's relevant legal requirements refer to as "student growth objectives" (SGOs).¹⁰ The MAP assessment includes on-line testing in reading and math in September/October (referred to herein as "pre-"), January (referred to as "winter" or "mid"), and April/May (referred to herein as "post-")¹¹; multi-format reporting of the results¹²; and the use of post-test target scores, referred to under the rubric of "RIT," based on the pre-test scores.¹³ Prior to the start of the school year, the District's office of evaluation used the NWEA PowerPoint to provide MAP training to the implementing administrators, including the elementary school principals and assistant principals. On a turnkey basis, the principals next provided the training to the teachers in their respective schools.

⁷ See *infra* note 52 and accompanying text.

⁸ The District used a different assessment for kindergarten and subsequently instituted an alternative at the high school level, but the MAP assessment applied to the grades that Chapman taught for the two years in question, 2014–15, and 2015–16. Conversely, the District's use of MAP in 2013–14 only applied to grades 6 and above, thus not pertinent to the instant case.

⁹ *Id.*

¹⁰ For the specific regulations, see *infra* note 56 and accompanying text.

¹¹ The reading and math tests are each 52 items on an adaptive basis, i.e., with the difficulty level adjusted of subsequent items adjusted based on the child's performance on the previous items.

¹² The format for the first report is a summary listing the pre-test, RIT, and post-test scores of all of the student's in the teacher's class.

¹³ This acronym is derived from the statistical concept of the **Rasch Unit** scale. See, e.g., <https://www.nwea.org/assessments/map/scale-and-norms/>

During the mandatory days before classes started in 2014–15, Forest Hills principal Daniel Corvi included training on the Danielson and MAP components for the school’s teaching staff. The contractually specified regular faculty meetings¹⁴ also included, along with various other administrative and instructional matters, follow-up training on these components of the mandatory evaluation process. For example, at the follow-up meeting on Oct. 6, 2014, which Chapman attended, he reviewed SGOs via the MAP system, including the access procedures for the summary and other reports at the NWEA dashboard for the MAP assessments.¹⁵ The final, “Updates Reminder” slide of the PowerPoint, which he customized from the turnkey NWEA training material,¹⁶ included the following statement: “SGO Draft Due: Tuesday, October 15th by midnight - based on RIT Projected Score.”¹⁷ The same slide also identified “Friday, October 31” as the final deadline.

Shortly before the Oct. 31 deadline, Colligan, on behalf of herself and Chapman, sought assistance from Corvi. He provided her with a sample SGO as an example.¹⁸ With Colligan in the lead, they developed SGOs for math and reading in a tiered format, grouping the students in

¹⁴ For 2014–15, these after-school faculty meetings were for 45 minutes on approximately a weekly basis, whereas for the following school year, they were changed to 90 minutes each approximately every two weeks.

¹⁵ E.g., Petitioner exhibit 10; Tr.-I at 192.

¹⁶ The incorporated NWEA slides included one that claimed that “SGOs are driven by teachers” rather than being “administrator-driven.” Petitioner exhibit 10, at 00816. Another NWEA slide identified as a misconception that “SGOs are a statistically precise measure of growth based on a pre-test/post-test model of performance.” *Id.* at 00818. The accompanying slide identifies some of the flaws of a pre-test/post-test model, including problems with the reliability of the results and the impossibility of extrapolating learning “from one data point.” *Id.* Instead, the slides claim that the MAP system for SGOs uses multiple measures for the starting point, and the illustrated format is a tiered approach, with 3–4 preparedness groups and with each group having a minimum target score that is “ambitious and achievable.” *Id.* at 00821-00822. The accompanying NWEA reference document only provides RIT means, or averages, for the beginning, middle, and end of each grade rather than individually determined RIT scores. *Id.* at 00265. Finally, the “frequently asked questions” sheet is quite clear that the SGO is worth 20% of the summative evaluation, but it is not at all reasonably clear that the target score will be based solely and individually on the child’s pre-test score. *Id.* at 00266.

¹⁷ *Id.*

¹⁸ Tr.-I at 376–77; Tr.-II at 447. The example was inferably tiered based on their result upon following its format. *Id.* at 378.

four tiers in successive ranges of MAP pre-test scores.¹⁹ The tiered format was similar to the aforementioned²⁰ illustrations. However, rather than being based on individual RIT projected scores, per the summary report available for reading and math, respectively, from the NWEA MAP dashboard,²¹ the target score for every tier, or group, was one point above the group's designated range.²² Additionally although incidentally, the number of students in each group was not completely accurate.²³

On Oct. 31, 2014, Chapman and Colligan submitted their two joint SGOs to Corvi. He signed the SGOs as approved after a quick immediate review.²⁴

During the rest of the 2014–15 school year, the three Chapman observations, including posthearing conferences, that Corvi and the assistant principal conducted are not at issue in this case.²⁵ The Jan. 12, 2015 training session, which both Chapman and Colligan attended, provided a sample SGO form that, although again using a tiered format, referred more specifically to RIT

¹⁹ Petitioner exhibits 11 and 13/Respondent exhibit 6.

²⁰ See *supra* notes 16 and 18.

²¹ Petitioner exhibit 16. This report includes for each student in the class the pre-test score along with its standard error (e.g., 173 and 3.0 for the first student on the list for math); the projected, or target, RIT score (e.g., 187 for this same student in reading) along with the corresponding growth projection and its standard error (14 and 4.2 for this student); and, after the post-testing, the post-test score and its standard error (173 and 2.9 for this student). The growth projection is based on the national sample of students at the same grade who participated in NWEA's MAP program. See, e.g., Respondent exhibit 33, at 00784.

²² Petitioner exhibit 11/Respondent exhibit 6. For both reading and math, not only the target but also the range for each group was uniform. For example, for the "medium" group in math, the selected range was 175–184, and the target score was 185. *Id.* at 0045–0046.

²³ For reading, the SGO identified the sizes of the first and third groups as 7 and 5 respectively, whereas the actual number of students in the SGO's specified ranges of pre-test scores was 8 and 6. For math, the SGO identified the sizes of the first and second groups as 9 and 6, whereas the actual sizes were 8 and 7. *Id.*

²⁴ *Id.*; Tr.-I at 380.

²⁵ The parties stipulated that each of these observations "substantially complied with evaluation procedures, does not contain any mistakes of fact, and is otherwise arbitrary or capricious." Joint exhibit 12.

projected, or target, scores.²⁶

On Saturday, May 2, 2015, Colligan e-mailed the joint SGO results to Corvi, “I just realized that Ms. Chapman and I never handed in our SGO spreadsheet.”²⁷ Early the following week when Colligan was at the principal’s office to pick up her mail, Corvi notified her that the results she had submitted were not in compliance with District requirements and had to be revised based on each student’s RIT target score in reading and math, respectively, as available on the testing reports accessible via the NWEA dashboard since soon after the MAP pre-testing. Revision of their SGOs based on the individual RIT target scores resulted in much lower overall summative SGO scores.²⁸ Corvi checked with District central officer administrators, who directed him to use the revised SGOs for the summative evaluation.²⁹ As a result, based on the revised scores, Corvi conducted a summative conference with Chapman on June 8, 2015, which he had postponed from the originally scheduled date of June 1, 2015 upon her request. The summary report showed the specific scores for the practice part (i.e., for the four Danielson domains and their respective components) and—per the aforementioned revision—for the

²⁶ Petitioner exhibit 18. Specifically, for example, the form for reading repeated the following SGO: “At least 70–84% of students in each preparedness group will achieve at least 1 year’s growth according to the individual projected MAP Language RIT Scale Norm.” *Id.* at 00250–00251. Yet, the use of multiple measures for each preparedness group (here, NJASK and DRA in addition to MAP) is illusory and confusing, because (1) only the MAP score counts for this growth calculation, and (2) the score ranges of the other two measures do not at all necessarily align with the specified range for MAP.

²⁷ Petitioner exhibit 44/Respondent exhibit 8. Again, their joint report was not carefully accurate. In reading, the group 1 results omitted the data for the student who had a pre-test score of 149 and a post-test score of 131; the group 2 results included a student who should not have been listed due to the absence of any pre-test score; the group 3 results listed post-test scores that were erroneously high for two students (181 and 175 rather than 177 and 172, respectively); and the single student in group 4 identified the post-test score as 201 whereas it was actually 188. The overall summative SGO score in reading should have been 3.55 rather than 3.70. *Id.* In math, the only error, although not materially affecting the summative SGO score, was again in the teachers’ direction; their listing of group 2 included a student who did not have a pre-test score but was counted as meeting their target. *Id.*

²⁸ Petitioner exhibits 12 and 14. The summative scores were 1.00 in reading and 1.00 in math. *Id.* If corrected for the errors attributable to the originally submitted SGOs (*supra* notes 19–23 and accompanying text), which were all in favor of the two teachers, these revised summative scores would remain the same because they were already at the lowest possible level on the four-point scale.

²⁹ Tr.-II at 226–27.

student learning part (i.e., the SGOs in reading and math) per their prescribed weights.³⁰ The resulting reported overall rating was Partially Effective.³¹ Consequently, he followed up at the same meeting with a corrective action plan (CAP).³²

In the summer of 2015, the District's turnkey evaluation training for principals included various revisions and refinements, such as adjusted domain weights for the teacher practice part and median student growth percentiles for the student learning part, but no notable clarification regarding the RIT format for SGOs.³³ The follow-up District training for principals in September 2015 noted inconsistencies in implementation of SGO's during the previous year.³⁴ To improve the process, the training provided additional information about SGOs, including a recommended the simple rather than tiered format (i.e., for the entire class rather than for subgroups of the class) for elementary teachers.³⁵ Yet, perhaps due to the introduction of an alternative assessment (which was not applicable to teachers on CAPs), this training did not specifically show the otherwise mandated individual RIT projection format.

During the 2015–16 school year, when Chapman had a grade 2 regular class of approximately 14 students, Corvi implemented the CAP,³⁶ and the District conducted the three requisite observations,³⁷ along with an additional one on May 25, 2016, at Chapman's request. Christie Whitzell, who was then senior lead teacher at the District's central office, conducted the

³⁰ The applicable weights were 20% for student learning and 80% for practice.

³¹ Petitioner exhibit 19/Respondent exhibit 9. The overall score was 2.53, and the range for Partially Effective was 1.85 to 2.64. *Id.*

³² Petitioner exhibit 20.

³³ Petitioner exhibit 54.

³⁴ Respondent exhibit 33, at 00782. For example, the PowerPoint noted: "Little guidance, clarity, and/or support for teachers across the district in creation, progress monitoring and/or scoring process." *Id.*

³⁵ *Id.* at 00790.

³⁶ E.g., Petitioner exhibits 21–22 and 38–40.

³⁷ E.g., Petitioner exhibits 23–28. Only the fourth observation was at issue; the other three were subject to a stipulation that they met the requisite standards within the arbitration context. *See supra* note 25.

additional observation.³⁸ In the interim, specifically on Feb. 1, 2016, Chapman and Corvi signed the SGO revision form based on a change in her class size.³⁹ Chapman’s SGO for this school year was in the simple rather than tiered format and based on individual RIT target scores, with the specified ranges for the four rating categories based on a District-wide mandated minimum of 75% rather than 70% for full, as compared with partial, effectiveness.⁴⁰

At the end of the year, Chapman’s calculated SGO level was 71% for reading and 71% for math, thus fitting in the Partially Effective range of 60%–74%.⁴¹ The resulting summative evaluation report, which included this SGO result for its designated 20% weighting and the practice domain scores for their designated weights within the remaining 80%, showed an overall score of 2.62, which was in the Partially Effective range (1.85–2.64).⁴²

As a result of the two consecutive annual summative ratings of Partially Effective, the District superintendent filed charges for termination based on inefficiency. The charges were duly processed through the requisite steps to this arbitration.⁴³

³⁸ Petitioner exhibit 29. She conducted the post-observation conference on June 6, 2015. See, e.g., Petitioner exhibits 30–31. For Chapman’s response, see Respondent exhibits 16–17.

³⁹ Respondent exhibit 21.

⁴⁰ *Id.*

⁴¹ *Id.* at 0225 and 0235. The format of the summary report changed slightly to show the range of RIT and percentile scores based on the standard error of measurement. *Id.* at 0227 and 0239. Moreover, this SGO report, unlike that for the previous year, showed that the RIT target scores for the two students who did not take the pre-test were based on their “winter” (i.e., mid-year) MAP testing results. *Id.* at 0229–0230.

⁴² Petitioner exhibit 45.

⁴³ E.g., Joint exhibits 1–4.

Applicable Statutes and Regulations

The TEACHNJ Act⁴⁴ includes various relevant requirements, including, for example:

- an evaluation “that incorporates analysis of multiple measures of student progress and multiple data sources”⁴⁵
- annual summative efficiency ratings of “ineffective,” “partially ineffective,” “effective,” and “highly effective”⁴⁶
- a “corrective action plan”⁴⁷
- a “school improvement panel”⁴⁸
- a Commissioner-approved “evaluation rubric”⁴⁹ that must include, *inter alia*, (1) the four aforementioned⁵⁰ summative rating categories⁵¹; (2) “multiple measures of practice and student learning,” with standardized achievement as a required but not predominant factor⁵²; (3) “multiple observations during the school year,” including an ongoing process for consistency⁵³; and (4) a pilot implementation program by Jan. 31, 2013.⁵⁴

The regulations for TEACHNJ specific to student growth objectives (SGOs), which are one of the two alternative methods for measuring the student learning component,⁵⁵ include the following excerpted provisions:

- (a) Measures of student achievement shall be used to determine impact on student learning. The student achievement measure shall include the following components:
-
2. [SGO(s)], which shall be specific and measurable, based on available student learning data, aligned to New Jersey Student Learning Standards, and based on growth and/or achievement. . . .

⁴⁴ N.J. REV. STAT. § 18A:6-117 *et seq.*

⁴⁵ *Id.* § 18A:6-119. For the definition of “multiple measures of student learning,” see *id.*

⁴⁶ *Id.* §§ 18A:6-119 and 18A:28-5(b)(3). For the specification of the annual written performance conference and report, see N.J. ADMIN. CODE § 6A:10-2.4(c)–(g).

⁴⁷ N.J. REV. STAT. § 18A:6-119.

⁴⁸ *Id.* § 18A:6-120. For the role and responsibilities of the SCIP and the over-arching District Advisory Evaluation Committee (DAEC), see N.J. ADMIN. CODE § 6A:10-2.4(b).

⁴⁹ N.J. REV. STAT. §§ 18A:6-122–123.

⁵⁰ See *supra* text accompanying note 46.

⁵¹ N.J. REV. STAT. § 18A:6-123(b)(1).

⁵² *Id.* §§ 18A:6-123(b)(4) and 18A:6-123(b)(6).

⁵³ *Id.* §§ 18A:6-123(b)(7) and §§ 18A:6-123(b)(11). For the types and timing of observations, see N.J. ADMIN. CODE § 6A:10-1.2.

⁵⁴ N.J. REV. STAT. § 18A:6-123(d). Moreover, the act required the prompt subsequent adoption of regulations the standards for evaluation rubrics. *Id.* § 18A:6-124.

⁵⁵ N.J. ADMIN. CODE § 6A:10-4.1.

- (e) [SGOs] for teachers shall be developed and measured according to the following procedures:
1. The chief school administrator shall determine the number of required [SGOs] for teachers. . . . A teacher without a student growth percentile shall have at least two and a maximum of four [SGOs]. . . .
 3. Each teacher shall develop, in consultation with his or her supervisor or a principal's designee, each [SGO]. If the teacher does not agree with the [SGOs], the principal shall make the final determination.
 4. [SGOs] and the criteria for assessing teacher performance based on the objectives shall be determined, recorded, and retained by the teacher and his or her supervisor by October 31 of each academic year
 5. Adjustments to [SGOs] may be made by the teacher in consultation with his or her supervisor only when approved by the chief school administrator or designee. Adjustments shall be recorded in the teacher's personnel file on or before February 15.
 6. The teacher's designated supervisor shall calculate each teacher's SGO score. The teacher's SGO score, if available, shall be discussed at the teacher's annual summary conference and recorded in the teacher's personnel file. . . .⁵⁶

The companion regulations for teacher evaluations require each school district to follow specified training procedures that include, *inter alia*, the following:

4. Annually require each supervisor who will conduct observations for the purpose of evaluation of a teacher to complete two co-observations during the academic year.
 - i. Co-observers shall use the co-observation to promote accuracy in scoring, and to continually train themselves on the instrument. . . .
5. Chief school administrators shall annually certify to the Department that all supervisors of teaching staff members in the school district who are utilizing educator practice instruments have completed training on the instrument and its application and have demonstrated competency in applying the educator practice instruments.⁵⁷

These regulations also have various requirements for the evaluation observations, including the following:

- (c) Each teacher shall be observed . . . at least three time during each school year but not less than once during each semester. For all teachers, at least one of the required observations shall be announced and preceded by a pre-conference, and at least one of the required observations shall be unannounced. . . . The following additional requirements shall apply:
 6. The teacher shall submit his or her written objection(s) of the evaluation within 10 teaching staff member working days following the conference. The objection(s) shall be attached to each party's copy of the annual written performance report.⁵⁸

⁵⁶ *Id.* § 6A:10-4.2.

⁵⁷ *Id.* § 6A:10-2.2(b).

⁵⁸ *Id.* § 6A:10-2.4.

The related statutory provisions for the termination of tenured teachers include the following:

- a. [Tenured teachers] shall not be dismissed . . . except for inefficiency, incapacity, or conduct unbecoming such a teaching staff member of other just cause⁵⁹
- 2) if the employee is rated partially effective in two consecutive annual summative evaluations . . . , the superintendent shall promptly file with the secretary of the board of education a charge of inefficiency, except that the superintendent upon a written finding of exceptional circumstances may defer the filing of tenure charges until after the next annual summative evaluation. . . .⁶⁰

The related applicable arbitration provisions provides the following considerations for the arbitrator in rendering a decision:

- a. [In an inefficiency case] in rendering a decision the arbitrator shall only consider whether or not:
 - (1) the employee's evaluation failed to adhere substantially to the evaluation process, including, but not limited to providing a corrective action plan;
 - (2) there is a mistake of fact in the evaluation;
 - (3) the charges would not have been brought but for considerations of political affiliation, nepotism, union activity, discrimination as prohibited by State or federal law, or other conduct prohibited by State or federal law; or
 - (4) the district's actions were arbitrary and capricious.
- b. In the event that the employee is able to demonstrate that any of the provisions of paragraphs (1) through (4) of subsection a. of this section are applicable, the arbitrator shall then determine if that fact materially affected the outcome of the evaluation. If the arbitrator determines that it did not materially affect the outcome of the evaluation, the arbitrator shall render a decision in favor of the board and the employee shall be dismissed.
- c. The evaluator's determination as to the quality of an employee's classroom performance shall not be subject to an arbitrator's review.
- d. The board of education shall have the ultimate burden of demonstrating to the arbitrator that the statutory criteria for tenure charges have been met.
- e. The hearing shall be held before the arbitrator within 45 days of the assignment of the arbitrator to the case. The arbitrator shall render a written decision within 45 days of the start of the hearing.⁶¹

⁵⁹ N.J. REV. STAT. § 18A:28-5.

⁶⁰ *Id.* § 18A:6-17.3(a).

⁶¹ *Id.* § 18A:6-17.2.

Discussion

Rather than separately reciting the parties’ arguments, I shall address them based on the foregoing factual findings and applicable legal standards under headings that correspond to the three challenges in the Respondent’s posthearing brief.⁶² I address the respective challenges in ascending order of their arbitrarily assessed strength. As an important contextual matter, this case arose during the transitional period of this state-operated district’s implementation of TEACHNJ.⁶³

Exceptional Circumstances⁶⁴

The Respondent claims that “the District should have found exceptional circumstances and granted her a third year.”⁶⁵ The statutory basis is the over-arching statutory allowance for districts to defer a tenure charge for inefficiency for an additional year after two consecutive summative evaluations of Partially Effective.⁶⁶

The Respondent argues that her particular pattern within the two consecutive summative

⁶² The Respondent’s prehearing arguments obviously were akin to a shotgun, but in light of the parties’ waiving of closing oral arguments, the prudently rifle-style after the evidentiary phase provide the controlling framework.

⁶³ The act, signed in August 2012, went into effect for the 2012–13 school year. http://www.njleg.state.nj.us/2012/Bills/PL12/26_PDF However, full implementation was pending a pilot program and regulations. *See supra* note 54 (N.J. REV. STAT §§ 18A:6-123 – 18A:6-124). Moreover, the District did not become state-operated until the end of 2012–13 and initiate its implementation until 2013–14. *See supra* notes 4–5 and accompanying text. Finally, the application of the District’s MAP-based student component to Chapman and similarly situated elementary teachers was not until 2014–15. *See supra* notes 8–10 and accompanying text.

⁶⁴ The Respondent’s posthearing brief (at 47) uses a broad “and/or” subheading for this challenge. As explained in this section, the primary focus is this particular statutory ground, and any cognizable alternate bases are subordinated and addressed within this section of the Opinion.

⁶⁵ The Respondent circularly argues that “the District, by and through [its superintendent], abused its discretionary powers and acted arbitrarily and capriciously by failing to investigate and find exceptional circumstances applicable to the within matter.” To the subordinated extent that the Respondent is raising the statutory ground of “arbitrary and capricious” conduct (*supra* note 61 and accompanying text: N.J. REV. STAT § 18A:6-17.2(a)), the arbitrator concludes that it has no separable successful application in this case. Assuming *arguendo* that there is a duty to investigate under the exceptional circumstances provision, the evidence in the case lacks the requisite trigger—whether it is reasonable suspicion or probable cause—to activate such a duty.

⁶⁶ *See supra* note 60 and accompanying text (N.J. REV. STAT § 18A:6-17.3(a)(2)).

evaluations of Partially Effective constitutes exceptional circumstances. More specifically she points to (1) the overall practice scores of Effective for both years, (2) the increase in the overall SGO scores, and (3) and the resulting overall summative score for the second of the two year being only .03 below the requisite minimum for the Effective category. This argument fails for two successive but independent reasons. First, such calculations are not exceptional, instead being part and parcel of the overall process, which includes a whole host of carefully weighted factors specified in state law and District policy. Second, even if this particular pattern did qualify as an exceptional circumstance, the applicable law specifically delegates the matter—via the clear statutory language of “may defer”⁶⁷—to school district discretion. Certainly in the context of a large state-operated district, this strict emphasis on teacher accountability and student achievement, does not constitute an abuse of that discretion.

Observational Violations

The Respondent targets three alleged fatal observer deficiencies in its second challenge: (1) the various observers’ alleged violation of the requirement for annual training/competency; (2) the observers’ alleged violation of the requirement to complete two observations annually; and (3) Whitzell’s failure to review Chapman’s observation rebuttal.

First, the aforementioned⁶⁸ regulatory requirement is for annual reporting that the supervisor-observers have completed training and demonstrated competency with regard to the teacher-practice instruments. Even if the regulation is interpreted to require annual training and competency, the evidence is sufficient that Chapman’s observers received regular training and

⁶⁷ *Id.*

⁶⁸ *See supra* note 57 and accompanying text (N.J. ADMIN. CODE § 6A:10-2.2(b)(5)).

exhibited competency in the Danielson instrumentation.⁶⁹ It is clear that proof does not come close to preponderant to fulfill one or more of the limited arbitral standards, such as arbitrary and capricious conduct, much less a resulting material effect on the Respondent's evaluation.⁷⁰

Second, the aforementioned⁷¹ regulatory requirement of two annual co-observations is expressly for the purpose of accuracy and training in scoring. Although the evidence is preponderant that none of the observers met this requirement in 2013-14,⁷² for the two years at issue for Chapman's evaluation the only identified violation is for Whitzell; her co-observations were limited to one in 2014–15.⁷³ On balance, the evidence is preponderant that this violation did not materially affect the outcome of Chapman's 2015–16 evaluation, which was the year in which she conducted an additional, final observation at Chapman's request. Although her observation scores were lower than the other observers, the difference was relatively modest,⁷⁴ and this difference may be reasonably attributable to other factors, such as the acceptable range for inter-observer agreement,⁷⁵ Whitzell's district-wide perspective, and this observation's

⁶⁹ The only evidence that Respondent identifies that arguably was to the contrary was the one-time completion of the Danielson proficiency exam. However, this exam was obviously for the initial threshold to perform observations, and competing it successfully at the end of the 2012–13 school year (e.g., Tr.-I at 48) is timely for purposes of conducting the observations at issue. The regular training of supervisors and administrators starting in December 2012 (Tr.-I at 23) and continuing the summers of the pertinent period (e.g., Tr.-II at 164) was certainly sufficient in the absence of any proof of nonproficiency.

⁷⁰ See *supra* note 61 and accompanying text (N.J. REV. STAT § 18A:6-17.2(a)).

⁷¹ See *supra* note 57 and accompanying text (N.J. ADMIN. CODE § 6A:10-2.2(b)(4)).

⁷² E.g., Tr.-I at 88; Tr.-II at 166.

⁷³ Tr.-I at 87–88. There is no evidence that the relationship of her scores with those of the co-observer indicated any inaccuracy or other lack of training, an area that the Respondent could have explored during the cross-examination.

⁷⁴ For the 1.0 to 4.0 scale, her domain 1 score was .1 less than the lowest other observer; for domain 2, the corresponding difference was .2; and for domain 3, it was .25. Petitioner exhibit 45. The principal was the major contributor for the final domain. E.g., Tr.-II at 240.

⁷⁵ Akin to the standard error of measurement for normed instruments such as MAP, inter-observer agreement is also known as interrater reliability, and its measurement does not have one consensus formula. E.g., KILEM LE GWET, HANDBOOK OF INTER-RATER RELIABILITY (2014).

unannounced status.⁷⁶ Additionally, she was a credible witness whose demeanor and testimony conveyed respectable professional expertise and diligence rather than any indication of evaluative bias or error with regard to the disputed observation. Finally, her extensive experience conducting the Danielson observations⁷⁷ and her reallocated responsibilities during 2015-16, upon her promotion to senior lead teacher,⁷⁸ mitigated the effect of her limited number of co-observations. Although this shortfall for the co-observation requirement is of concern and merits requisite attention, in this case it does not provide the basis for revising or reversing the summative evaluation for the second year, particularly in light of the statutory limitation of the arbitrator's authority in terms of the quality of an employee's classroom performance.⁷⁹

The final alleged observation violation pertains to the regulatory requirement for any written teacher objections after the post-observation conference.⁸⁰ The Respondent characterizes the objections as “rebuttals” and claims that Whitzell’s failure to review and respond to Chapman’s timely objections circumvented their “entire purpose.” However, the regulation does not express or imply the purpose of being a debate, requiring in effect a surrebuttal. Rather, it merely provides for a teacher to add her perceptions to those of the observer for review at the summative and any subsequent stages.⁸¹ Although it would have been best practice for Whitzell to review and respond to Chapman’s objections, the regulation does not require this action, and

⁷⁶ For the regulatory requirements for announced and unannounced observations, *see supra* note 58 (N.J. ADMIN. CODE § 6A:10-2.4(c)). In the District for the 2015–16 school year, unannounced observations were additionally different due to their length—a minimum of 20 minutes compared to a minimum of 40 minutes for the announced observations. Tr.-I at 33.

⁷⁷ She had performed well over 100 such observations since 2013-14. Tr.-I, at 39.

⁷⁸ She conducted less observations in 2015–16 because her position required more time with curriculum and professional development. *Id.* at 88.

⁷⁹ *See supra* note 61 and accompanying text (N.J. REV. STAT. § 18A:17.2(c)).

⁸⁰ *See supra* note 58 and accompanying text (N.J. ADMIN. CODE § 6A:10-2.4(c)(6))

⁸¹ There was no evidence that Corvi did not consider Chapman’s side of the story or that Chapman raised the issue at or before her summative conference for 2015–16. Moreover, the contents of her objections largely amount to a difference of professional perception as to the quality of her performance, which—again (*supra* note 79)—is beyond arbitrator’s authority.

its absence does not meet the applicable, restricted two-part standard,⁸² such as being not only arbitrary and capricious but also prejudicial to Chapman’s summative evaluation for 2015-16.

In sum, although the Respondent partly proved two of the three alleged observation deficiencies, neither one nor their combination met the requisite standard for reversal. They did not amount to substantial noncompliance, mistaken fact, or arbitrary and capricious conduct that materially affected the outcome of the Chapman’s evaluation for 2014–15 and/or 2015–16.⁸³

2014–15 SGO

The Respondent’s final challenge presents a close call.⁸⁴ The factual basis for the challenge is relatively simple: (1) relying on the principal’s training and assistance, Chapman and Colligan developed their joint SGOs for 2014–15 and submitted them to Corvi by the designated deadline in October; (2) Corvi promptly approved them; and (3) upon submission of the results of the SGOs in May, Corvi unilaterally rejected and revised them to scoring based on the students’ RIT projected scores.

The preliminary part of analysis is to eliminate two distractors, or red herrings. First, the Respondent argues that, under the applicable regulations, “only a teacher” may initiate changes to his or her SGO and that no such changes are allowable after February 15, especially without

⁸² See *supra* note 61 and accompanying text (N.J. REV. STAT § 18A:6-17.2(a)).

⁸³ *Id.* There was no allegation with regard to the other statutory standard, which concerns various civil rights or “other conduct prohibited by State or federal law.” *Id.*

⁸⁴ Although preponderant proof is indisputably applicable quantum, the evidence, however, is not so closely balanced as to be in “ equipoise.” See, e.g., *Schaffer ex rel. Schaffer v. Weast*, 546 U.S. 49, 58 (2005). Thus, burden of proof is not determinative for this issue, much less for the previous two issues. This situation is fortunate here because the TEACHNJ statute, although expressly putting the burden of persuasion on the District for the tenure charges, is not clear as to whether the teacher’s burden to “demonstrate” one or more of the four arbitral standards signifies burden of production, burden of persuasion, or burden shifting. See *supra* note 61 and accompanying text (N.J. REV. STAT. § 18A:6-17.2(b) and *id.* § 18A:6-17.2(d)).

any written memorialization of the amendment.”⁸⁵ Yet, the regulations expressly leave the final determination of the SGOs to the principal, and they establish a Feb. 15 deadline for adjustments “made by the teacher.”⁸⁶ Thus, the regulations do not prohibit the administration from correcting an improper SGO at any time during the school year, with the inferable proviso of sufficient justification.⁸⁷ Moreover, notice is not the key in this case because Chapman received effective notice inasmuch as Colligan, as her delegated intermediary, immediately communicated the situation to her.

Second, in light of the commissioner’s approval,⁸⁸ the Respondent’s broad-based contention about the legality of the District’s MAP assessment is, like any corresponding challenge to TEACHNJ, belongs in another forum. More specifically, the extent that the Respondent’s challenge extends on a system-wide policy basis to TEACHNJ as applied via the District’s use of MAP RIT projected scores, this issue, like the fairness or legality of TEACHNJ on its face, is either for the state legislature or the courts, which have been notably deferential to the defendants.⁸⁹

Instead, a careful review of the evidence in relation to the applicable legislation and

⁸⁵ Respondent’s posthearing brief at 33. For the notice, Respondent alternatively relies on basic notions of fairness and due process, citing confirmation in the testimony from the District’s evaluation director. *Id.* Yet, the evaluation director’s testimony neither supported the limitation to written notice nor the limitation to Feb. 15. Tr.-I at 233–43.

⁸⁶ *See supra* note 56 and accompanying text (N.J. ADMIN. CODE § 6A:10-4.2(e)(2) and *id.* § 6A:10-4.2(e)(5)). However, contrary to the Petitioner’s posthearing brief (at 14), the chief school administrator only determines the number of the SGOs. *Id.* § 6A:10-4.2(e)(1).

⁸⁷ In the absence of express specifications, this determination, as shown *infra*, ultimately depends on the circumstances, with due weight for objective and equitable considerations.

⁸⁸ *See supra* text accompanying notes 9–10. Although not finding it necessary to address this issue within the confines of this case, the arbitrator cannot help but wonder whether the MAP RIT projection system meets TEACHNJ’s repeated and varied references to “multiple measures” in relation to student growth. N.J. REV. STAT §§ 18A:6-119 et al. The statute allocates this determination to the commissioner on an annual basis. *Id.* §§ 18A:6-122 and 18A:6-122. For subsequent years and less narrowly confined cases, this issue, if not resolved more definitively and defensibly, may well become the key in future arbitral or judicial proceedings.

⁸⁹ *See, e.g.,* Cook v. Bennett, 792 F.3d 1294 (11th Cir. 2015); Wagner v. Haslam, 112 F. Supp. 3d 673 (M.D. Tenn. 2015); State *ex rel.* Stapleton v. Skandera, 346 P.3d 1191 (N.M. 2015).

regulations reveals that resolving the SGO issue, within the ambit of the arbitral standards,⁹⁰ warrants in this case a two-step, flowchart-type objective analysis.⁹¹ The first step is to determine the expectations for 2014–15 SGOs that a reasonable teacher would understand under the circumstances at Chapman’s time and place. The second step is to determine whether Chapman’s 2014–15 SGO fit within the reasonable range of those expectations.

For the first step, the conclusion is inescapable that the NWA turnkey training materials upon which the District relied were confusing and conflicting. The MAP system that projected RIT target scores for the post-testing based solely on pre-test scores came closer to the “misconceptions” identified in NWA’s training slides⁹² than the content of this training.⁹³ For example, the emphasis on SGOs being teacher-, rather than administrator-, driven lacks the nuance in the reality of state law and District policy.⁹⁴ Finally, the examples in the training materials and the principal’s assistance were tiered,⁹⁵ thus interfering with the ultimate priority on individual RIT projected scores.⁹⁶ In contrast, the evidence about the other teachers at Chapman’s school was a “wash,” thus having no significant effect on reasonableness other than

⁹⁰ See *supra* note 61 and accompanying text (N.J. REV. STAT. § 18A:6-17.2(b)). Contrary to the Petitioner’s posthearing brief (at 16), if Chapman passed this two-step test, the material effect would be use of her unchanged SGOs, thus resulting in an Effective summative rating for 2014–15.

⁹¹ This objective approach keys on reasonableness under the circumstances.

⁹² See *supra* note 16.

⁹³

⁹⁴ See *supra* text accompanying note 86.

⁹⁵ See *supra* notes 16 and 18.

⁹⁶ More specifically, while the final slide referred to the required SGO basis being RIT projected scores (*supra* text accompanying note 17), the preceding content referred to multiple initial sources rather than the MAP pre-test score for the baseline and showed unexplained averages per tier rather than the individual RIT projected scores (*supra* note 16). What is sorely missing for that year and not part of the training until 2015–16 (*supra* text accompanying note 35) is the basic template of a simple (i.e., nontiered) example showing the pre-test and, as otherwise available in the initial summary report, the projected MAP scores on an individual student-by-student basis.

failing to confirm the necessity of a narrow uniform understanding.⁹⁷ Thus, a reasonable range of expectations for a teacher in Chapman’s circumstances, including the aforementioned⁹⁸ context of a transitional year with acknowledged inconsistencies,⁹⁹ was to produce SGOs with a tiered (or simple) format and based on RIT projected scores or target scores that were otherwise ambitious yet achievable.¹⁰⁰

However, at the second step, rather than the purported formulation of the target scores based on Colligan’s and Chapman’s assessment of multiple sources after a month and a half of student data,¹⁰¹ including tiers based on similar considerations,¹⁰² the joint SGOs in reading and math were uniformly based solely on the MAP pre-test score, and the tiers for both SGOs were uniformly based on 10-point intervals.¹⁰³ Moreover, the target scores were uniformly and rather meagerly one point above the designated range for each tier.¹⁰⁴ As a result, their projected scores represented average gains of 5.3 points and 6.3 points and for reading and math, respectively,

⁹⁷ On the one hand, Colligan testified that two other pairs of co-teachers at the school had submitted SGOs that were not based on RIT projected scores and that the Principal had similarly changed two of them. Tr.-I at 406–10. However, her testimony was generally not entirely credible, being replete with one-sided overstatement. Moreover, the Chapman’s response to the charges and various prehearing discovery documents, including the required “list of witnesses with a complete summary of their testimony” (N.J. REV. STAT. § 18A:6A-17.1), did not provide any warning about this evidence or request for corroborating SGOs (e.g., Joint Exhibits 7–12), and the Respondent did not provide testimony or documentation from any of these four teachers. Completing the neutralization of this factor, the principal rather vaguely testified that he had no recollection of receiving and changing such SGOs. Tr.-II at 317–19. When the arbitrator tried to elicit clarification, including whether he his recollection was definitive or, instead, he did not recall either way, he ambiguously replied: “not so even to my vivid recollection.” *Id.* at 319.

⁹⁸ *See supra* note 63 and accompanying text.

⁹⁹ *See supra* note 34 and accompanying text.

¹⁰⁰ This overall standard was reasonable inferable in the focus of TEACHNJ on “improved student outcomes, including objective measures of student growth” (N.J. REV. STAT. § 18A:6-118(b)); the District’s state-operated particular priority on raising student achievement (*see supra* text accompanying note 4); and the training materials accurate use of these reasonably rigorous terms (*supra* note 16).

¹⁰¹ Tr.-I at 377.

¹⁰² Tr.-II at 338.

¹⁰³ *See supra* text accompanying note 19. The use of the pre-test MAP scores also contradicts Colligan’s testimony that she could not access the RIT data. Tr.-I at 446. Colligan’s assertions apply as well to Chapman based on her proxy-like reliance on Colligan for the formulation and communication of the SGOs during 2014–15. E.g., Tr.-II at 337 and 388–89.

¹⁰⁴ *See supra* text accompanying notes 20–21.

compared to the RIT projected scores corresponding average gains of 14.3 and 15.2 points.¹⁰⁵

The inescapable conclusion is that the targets in Chapman’s 2014–15 SGOs, which amounted to well under half of the generally applicable targets, were much too achievable but far less than ambitious, thus not at all fitting in the reasonably expected range. For Chapman to rely on the principal’s cursory acceptance of her SGOs, akin to equitable estoppel, is not sustainable in light of her equal lack of diligence.¹⁰⁶

Finally, the arbitrator finds that neither the District’s nor the teacher’s actions within this two-step analysis were intentional or in bad faith. Rather akin to negligence, under either the standard of arbitrary and capricious conduct or that of mistake in fact,¹⁰⁷ Chapman ultimately negated her claim with a commensurate contribution via her lack of due care in implementing her part of the accountability process. As a matter of the applicable law, with due weight to considerations of equity, this close call cannot be fairly in her favor. The arbitrator is without authority to be an angel of mercy or to apply his views of educational effectiveness and accountability to the MAP system and the District’s administration of it.

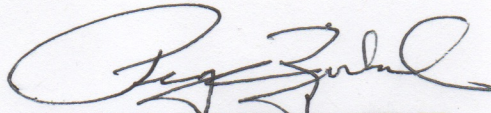
¹⁰⁵ The arbitration calculated these results by (1) applying the target scores in the two SGOs (Petitioner’s exhibits 11 and 13) to the students’ pre-test scores in reading and math for 2014–15 (Petitioner’s exhibit 16) and computing the corresponding averages of the “SP 15 Growth Projection” column in the same 2014–15 MAP summary report (*id.*). Each entry in this column represents the difference between the pre-test score and the projected RIT score.

¹⁰⁶ Reflecting the lack of care, which further effectively lowered their SGO bar, the calculations included various errors. *See supra* notes 23 and 27.

¹⁰⁷ For the elimination of the other alternative arbitral standard that the Respondent raised, which is substantial adherence standard to the evaluation process, *see supra* notes 85–89 and accompanying text.

Award

In sum, after careful attention to the parties' arguments, the applicable law, and the case evidence, the arbitrator concludes that the District's tenure charges for inefficiency are sustained. The dismissal of Gerri A. Chapman is upheld.



Perry A. Zirkel, Arbitrator

12/13/16
Date

COMMONWEALTH OF PENNSYLVANIA
NOTARIAL SEAL
ANGELA MARIE CARWELL
Notary Public
SOUTH WHITEHALL TWP., LEHIGH COUNTY
My Commission Expires Jan 27, 2017

