TEACHER-SCHOOL BOARD GRIEVANCE ARBITRATION AWARDS IN THE COURTS: FACTS AND FIGURES

*Perry A. Zirkel**

Collective bargaining is authorized for public school teachers in approximately 35 states, with the remaining state laws either silent or prohibitive.[1] The majority of these 35 states include grievance procedures as a mandatory subject of bargaining.[2] Yet, despite its importance as the culminating, binding, and third-party step in the grievance process, teacher-board arbitration has received only limited empirical attention.[3]

A particular gap concerns judicial review of the resulting arbitration awards. For example, what is the frequency and location of court decisions that review teacher-board arbitration awards? Do these judicial rulings support the general view that courts almost always uphold arbitrators under doctrines of deference and finality? And do the judicial outcomes differ for the arbitral rulings concerning arbitrability from those concerning the vacatur of the award?

The purpose of this article is to address these questions via an empirical analysis of the pertinent case law. The frame of reference for the analysis consists of the legal backdrop for labor arbitration and the empirical backdrop of prior research, each largely in the successively surrounding sectors for teacher-board arbitration.

---

* **Perry A. Zirkel** is University Professor Emeritus of Education and Law at Lehigh University. His 40-year career as an attorney in various neutral roles includes serving as a part-time labor arbitrator with membership in the National Academy of Arbitrators.

[1] See, e.g., Emily Workman, State Collective Bargaining Policies for Teachers (Dec. 2011), http://www.ecs.org/clearinghouse/99/78/9978.pdf (last visited Dec. 28, 2015). The number and scope of these laws are subject to flux, as illustrated in Wisconsin's relatively recent political controversy. Teachers, Inc. v. Walker, 851 N.W.2d 337 (Wis. 2014) (upholding constitutionality of 2011 amendments that added various significant restrictions on state's public employee collective bargaining law).

[2] See, e.g., National Council on Teacher Quality, State Influence: Grievance Procedures (n.d.), http://www.nctq.org/districtPolicy/stateInfluence.do (last visited Dec. 28, 2015).

[3] See, e.g., Frederick M. Hess & Andrew P. Kelly, *Scapegoat, Albatross or What?: The Status Quo in Teacher Collective Bargaining, in* COLLECTIVE BARGAINING IN EDUCATION 53, 85 (Jane Hannaway & Andrew J. Rotherham eds., 2006) (observing that "[g]rievance arbitration is a quasilegal, poorly understood process [that despite its importance] . . . . has largely escaped . . . scholarly . . . attention").

I. LEGAL BACKDROP

The applicable posture and standards for judicial review of teacher-board grievance arbitration awards is the result of three successive legal frameworks. The first two, as explained in more detail elsewhere,[4] are the Federal Arbitration Act (FAA) of 1925 and the *Steelworkers* Trilogy of Supreme Court decisions in 1960.[5] The third and culminating framework consists of the state laws and court decisions specific to judicial review in the teacher-board context of grievance arbitration.

A. Federal Arbitration Act

Originally intended primarily for the commercial and maritime contexts,[6] the FAA established the framework for a broad-based judicial receptivity to grievance arbitration.[7] In addition to establishing the enforceability of written agreements for arbitration,[8] the Act authorizes judicial vacatur[9] for limited reasons largely concerning the arbitral process.[10] The only one specific to the product—the "award," or written arbitration decision—focuses on the alternatives of

---

[4] See, e.g., Michael H. LeRoy & Peter Feuille, *Happily Never After: When Final and Binding Arbitration Has No Fairy Tale Ending*, 13 HARV. NEGOT. L. REV. 167, 173–82 (2008).

[5] FAA commercial labor arbitration and Trilogy unionized labor arbitration are fundamentally different. See, e.g., Mitchell H. Rubinstein, *Altering Judicial Review of Labor Arbitration Awards*, 2006 MICH. ST. L. REV. 235, 242. Nevertheless, they overlap and interact. See, e.g., Mark Berger, *Arbitration and Arbitrability: Toward an Expectation Model*, 56 BAYLOR L. REV. 753, 759 (2004).

[6] See, e.g., Michael H. LeRoy, *Irreconcilable Differences?: The Troubled Marriage of Judicial Review Standards under the Steelworkers Trilogy and the Federal Arbitration Act*, 2010 J. DISP. RESOL. 89, 98 (2010).

[7] 9 U.S.C. §§ 1–16 (2013).

[8] *Id*. The Supreme Court has interpreted this provision as reflecting a policy under the FAA in favor of arbitrability. Moses H. Cone Mem'l Hosp. v. Mercury Constr. Co., 460 U.S. 1, 24–25 (1983).

[9] Specifically, the Act also authorizes the court, upon either party's motion, to "modify" or "correct" arbitral awards in specified, limited circumstances. *Id*. § 11. Conversely, the Act refers to a court's authority, upon a party's motion, to "confirm" an award. *Id*. § 13.

[10] *Id*. § 6: "(1) where the award was procured by corruption, fraud, or undue means; (2) where there was evident partiality or corruption by the arbitrators, or either of them; (3) where the arbitrators were guilty of misconduct in refusing to postpone the hearing, upon sufficient cause shown, or in refusing to hear evidence pertinent and material to the controversy; or of any other misbehavior by which the rights of any party have been prejudiced."

exceeding or imperfectly executing arbitral authority.[11] Moreover, the express exclusion for "the merits,"[12] along with the Act's legislative history,[13] reflect an intent for restrictive judicial review.[14]

B. Steelworkers Trilogy

Targeting the collective bargaining context generally referred to as labor rather than commercial arbitration and doing so primarily as a matter of common rather than statutory law,[15] the Supreme Court issued three companion decisions 35 years after the passage of the FAA that similarly provided for deferential judicial review. In the Trilogy, the Court specifically addressed the separable, but overlapping, issues of arbitrability and vacatur. Providing a broad presumption in favor of substantive arbitrability,[16] the Court prescribed a "positive assurance" standard in *United Steelworkers v. Warrior & Gulf Navigation Company*.[17] Conversely providing a

[11] *Id*.: "(4) where the arbitrators exceeded their powers, or so imperfectly executed them that a mutual, final, and definite award upon the subject matter submitted was not made."

[12] *Id*. § 11(b)–(c).

[13] Arbitration of Interstate Commercial Disputes: Joint Hearings on S. 1005 and H.R. 646 Before the Subcomm. of the Comms. on the Judiciary, 68th Cong. 36 (1924) ("There is no authority and no opportunity for the court, in connection with the award, to inject its own ideas of what the award should have been.")

[14] For the pro- and anti-expansion judicial interpretations of the statutory and common law grounds for vacatur under the FAA, see Thomas S. Meriwether, *Limiting Judicial Review of Arbitration Awards under the Federal Arbitration Act: Striking the Right Balance*, 44 HOUS. L. REV. 739, 750–58 (2007). For a subsequent decision expressly limited to the FAA, see Hall Street Assoc., L.L.C. v. Mattel, Inc., 552 U.S. 576, 584 (2008) (ruling that the FAA's specified grounds for vacatur are exclusive).

[15] The underlying federal statute was the Labor Management Relations Act of 1947, ch. 120, § 301, 61 Stat. 136, 156-57 (1947) (current version at 29 U.S.C. § 185(b) (2013)). Setting the stage for the Trilogy, the Supreme Court authorized the development of such common law under § 301 of this statute in *Textile Workers v. Lincoln Mills*, 353 U.S. 448 (1957). For an overview and analysis of foundational Supreme Court decisions concerning arbitration under § 301 subsequent to the Trilogy, see, e.g., George W. Moss, *The Fate of Arbitration in the Supreme Court: An Examination*, 9 LOY. U. L.J. 369 (2015).

[16] In contrast, the Trilogy did not specifically address procedural arbitrability. Not long thereafter, however, the Court largely reserved this matter for the arbitrator's initial determination. John Wiley & Sons v. Livingston, 376 U.S. 543, 557 (1964) ("Once it is determined . . . that the parties are obligated to submit the subject matter of a dispute to arbitration, 'procedural' questions which grow out of the dispute and bear on its final disposition should be left to the arbitrator.").

[17] United Steelworkers of Am. v. Warrior & Gulf Navigation Co., 363 U.S. 574, 582–83 (1960) ("An order to arbitrate the particular grievance should not be denied unless it may

restrictive posture for vacatur, the Court established, in *United Steelworkers v. Enterprise Wheel & Car Corporation*, an "essence" test[18] that is unmistakably deferential.[19] The remaining decision in the Trilogy, while focused on the threshold arbitrability issue, served as an over-arching reminder of judicial deference to what the parties had collectively bargained.[20]

C. Teacher-Board Context

Inasmuch as teacher-board grievance arbitration is a matter of public employees collectively bargaining under state law rather than private employees under individual or collective contracts under federal law, the foregoing two frameworks serve only indirectly for the specifically applicable judicial standards for arbitrability and vacatur. The intervening development was the Uniform Arbitration Act (UAA), generally resulting—with very limited differences—in state statutory standards for vacatur that were substantially the same as those under the FAA.[21] However, the courts in an increasing but still limited number of states have developed a variety of additional standards for vacatur, principally being (1) manifest disregard of the law[22] and (2) the public policy exception.[23] Finally, the panoply of

---

be said with positive assurance that the arbitration clause is not susceptible of an interpretation that covers the asserted dispute").

[18] United Steelworkers of Am. v. Enterprise Wheel & Car Corp., 363 US. 593, 597 (1960) ("[a labor arbitrator's award] is legitimate only so long as it draws its essence from the collective bargaining agreement"). The "only" in this context cross-refers to the overriding contractual boundary, such that "an arbitrator is confined to interpretation and application of the collective bargaining agreement; he does not sit to dispense his own brand of industrial justice." *Id.*

[19] *Id*. at 599 ("the courts have no business overruling [the arbitrator] because their interpretation of the contract is different from his").

[20] United Steelworkers of Am. v. Am. Mfg. Co., 363 US. 564, 597 (1960) ("the moving party should not be deprived of the arbitrator's judgment, when it was his judgment and all that it connotes that was bargained for").

[21] See, e.g., Stephen K. Huber, *State Regulation of Arbitration Proceedings: Judicial Review of Arbitration Awards by State Courts,* 10 CARDOZO ONLINE J. CONFLICT RESOL. 509, 521–22 (2009); Stephen Wills Murphy*, Judicial Review of Arbitration Awards under State Law*, 96 VA. L. REV. 887, 891–92 (2010).

[22] See, e.g., Murphy*, supra* note 21, at 911–12. For the origin of this standard, see *Wilko v. Swan*, 346 U.S. 427, 436–37 (1953). For more recent refinement, see, e.g., *Eastern Coal Corp. v. United Mine Workers of Am*., 531 U.S. 57, 62–63 (2000).

[23] See, e.g., Murphy*, supra* note 21, at 911–12. For the origin of this standard, see *W.R. Grace & Co. v. Local 759*, 461 U.S. 757, 766 (1983).

statutory and case law applicable to teachers and school boards provide an intersecting overlay affecting arbitrability and vacatur.[24]

II. EMPIRICAL BACKDROP

Empirical research specific to grievance arbitration[25] is largely limited to the FAA and the Trilogy contexts, rather than the teacher-board setting. The next two sections review the recent research literature at the successive levels of grievance arbitration and judicial review, with the primary focus on outcomes. The coverage does not extend to analyses published prior to the most recent decade, except to fill notable gaps in relevant findings during this period.

A. Arbitration Level

Relatively recent published analyses of outcomes of grievance arbitration cases generically, i.e., those not specific to the teacher-board context, are relatively extensive.[26] However, they are largely limited in several respects, including the issues,[27] the outcome scale,[28] the jurisdiction,[29] the sector,[30] and the source(s).[31] These variations

[24] This varying panoply includes statutory and case law applicable to (a) teacher-board collective bargaining, (b) teacher status (e.g., certification, evaluation, nonrenewal, and termination); and, to a lesser extent, (c) federal requirements (e.g., No Child Left Behind Act and Family and Medical Leave Act).

[25] The focus here is on traditional grievance arbitration, that is, in the unionized context. Thus, this review does not extend to the extensive literature specific to (a) grievance arbitration for nonunion employees, which is often referred as "mandatory arbitration" or "employment arbitration"; (b) interest arbitration; or (c) advisory arbitration.

[26] In contrast, empirical analyses of arbitration awards specific to arbitrability are infrequent and not recent. See, e.g., Perry A. Zirkel, *Procedural Arbitrability of Grievance Cases*, 13 J. COLLECTIVE NEGOT. 351 (1984) (finding that 85% of a sample of published arbitration awards during the period 1978–81 rejected claims of procedural nonarbitrability, with 68% of the cases being in the private sector and timing being the most frequent procedural issue).

[27] For example, the issue-specific analyses are often limited to discipline, especially discharge, cases.

[28] For example, the usual metric is a two-category win-loss scale that fails to differentiate inconclusive as well as mixed conclusive outcomes. The inconclusive category is more typical of arbitral outcomes upon judicial review, such as denial of summary judgment or remand to the arbitrator for further consideration. However, mixed outcomes are not unusual at the arbitral level for cases that have more than one issue and/or for awards that are partially in favor of each side.

[29] For example, these analyses are often limited to a single state.

limit generalizability not only for each analysis but also for them collectively.

A sample of such recent analyses, in overlapping chronological order is illustrative of this variance. First, based on a sample of 175 published awards specific to workplace violence in both the public and private sectors for various selected years between 1979 and 2001 and the traditional win-loss categorization, Gely and Chandler found the following outcomes distribution: Union wins – 54% and Union losses – 46%.[32] Serving as a second example, based on a sample of 200 awards published in *Labor Arbitration Reports* for the period 1988–2003 specific to discipline in the public sector, LaVan found the following outcomes distribution: in favor of grievant – 42%, "split" – 21%, in favor of employer – 38%.[33] Third, Bognanno et al. found the following outcomes distribution of grievance arbitration awards for a large sample "rough[ly]" representative of public and private sector discharge cases in Minnesota for the period 1982–2005: completely for grievant – 20%, partially for each side – 28%, and completely for employer – 52%.[34]

---

[30] For example, they range from a specific subsector, such as police or teacher cases in the public sector, to coverage of both the public and private sectors.

[31] For example, the coverage may be limited to cases "published" in one or more labor arbitration reporting services, such as the Bureau of National Affairs' LABOR ARBITRATION REPORTS, or based more broadly on cases collected in one of the labor arbitration appointment services, such as the AAA.

[32] Rafael Gely & Timothy D. Chandler, *Exploring the "Lumpiness" of Grievance Arbitration Decision Making*, 32 J. COLLECTIVE NEGOT. 287 (2008). For the outcomes categorization, they defined a win as those awards in which the arbitrator found in favor of the grievant in whole or in part and a loss as those in which the arbitrator upheld the employer's disciplinary action in full. *Id.* at 291. This attribution of compromise, or partial, outcomes as being Union wins suggests a less than objective viewpoint.

[33] Helen LaVan, *Arbitration of Discipline in the Public Sector: Case Characteristics and Party Behavior Predicting Case Outcome*, 31 J. COLLECTIVE NEGOT. 199, 204 (2007). The boundaries and representativeness of her case sampling as well as the terminology and use of her outcomes scale are subject to question.

[34] Mario F. Bognanno, Jonathan E. Booth, Thomas J. Norman, Laura J. Cooper, & Stephen F. Befort, *The Conventional Wisdom of Discharge Arbitration Outcomes and Remedies: Fact or Fiction?* 16 CARDOZO ONLINE J. CONFLICT RESOL. 153 (2014). Not providing for inconclusive outcomes, their four-category scale differentiated the partial awards into those providing no backpay (10.4%) and those providing no backpay (17.4%). However, these distinctions are limited to discharge cases and are questionable because they do not take into consideration the varying time lag between the discharge date and the award date.

The published research specific to arbitration in the K–12 education context is relatively scant, even when subject to a broader scope of review in terms of chronology and methodology. First, in an analysis of 244 awards published in CCH's LABOR ARBITRATION AWARDS for the period 1989–1998, Zirkel and Miller found the following overall outcomes distribution: in favor of the grievant – 43%, "compromise"-10%, and in favor of the school district – 47%.[35]

Second, in a qualitative study using a case study design in six school districts in Florida for the period 2004–2008, Osborne-Lampkin found that the arbitration step was the exception in school teacher grievances.[36] More specifically, she found that none of the districts had proceeded to arbitration in their resolution of teacher grievances during the five-year period of the study.[37] Possibly attributable to the absence of arbitration activity, she found that legal considerations, including the odds of judicial vacatur, played a relatively minor role among the various factors that contributed to the parties' choices as to when and how to resolve the grievance.

B. Judicial Level

For the framework of applicable standards of judicial review, Murphy canvassed and categorized state statutes, finding that 47 states and the District of Columbia have adopted legislation based on either the UAA (n=38) or the Federal Arbitration Act (n=9), which provides for vacatur based on specified arbitrator misconduct and where the award exceeded the arbitrator's authority. [38] He further reported that the courts in 38 of these jurisdictions have interpreted their statutes as precluding judicial review of arbitrators' factual findings, legal conclusions, or both. Conversely, courts allow for vacatur for "manifest disregard of law" in 18 of these jurisdictions

---

[35] Perry A. Zirkel & Chad C. Miller, *Grievance Arbitration in K–12 Education Cases: Do Selected Case Characteristics Make a Difference?*, 28 J. COLLECTIVE NEGOT. 295, 297 (1999). For the grievants, the scope extended beyond teachers, with 27% being clerical or other classified employees and 7% being paraprofessionals. For the outcomes scale, the intermediate category of "compromise" represented rulings partially upholding and partially denying the grievance.

[36] La'Tara Osborne-Lampkin, *Grievance and Arbitration Practices and Decisions in Schools: Outcomes of Rational Decision Making?* 20 J. SCH. LEADERSHIP 491 (2010).

[37] *Id*. at 516. According to her account, the Bureau of National Affairs had reported only two arbitration awards for the entire state during this period. *Id*. at 517.

[38] Stephen Wills Murphy, Comment, *Judicial Review of Arbitration Awards under State Law*, 96 VA. L. REV. 887 (2010).

and for the "public policy exception" in an unspecified but smaller number of these jurisdictions. However, his analysis was generically broad-based rather than specific to either labor arbitration or public school teacher grievances.

Within the framework of state or federal statutes, two relatively recent studies have analyzed the court decisions that have reviewed grievance arbitration awards beyond the specific context of teacher-board collective bargaining agreements. First, in an analysis limited to a leading state, Skanes analyzed 146 New York appellate court decisions that reviewed a wide variety of arbitration awards during the period 1999–2009, finding that the vacatur rate in the labor-employment cases was 30%.[39]

Second, in a national analysis of 281 federal court decisions during the period 2001–2006 that confirmed or vacated union-employer arbitration awards, LeRoy and Feuille found that the confirmation rates of the federal district and appellate courts were 77.6% and 76.3%, respectively.[40] These rates were slightly higher than those for their corresponding analyses for earlier periods, but their revisions in search and selection procedures precluded precise comparisons.[41] The most common issue was termination (46%),[42] and the most frequent legal basis was the essence test. However, their scope excluded state court decisions, thus most of the teacher-district arbitration cases, and their search process was skewed toward this Trilogy-related standard.

---

[39] Monica R. Skanes, Comment, *The Truth Behind "Final and Binding" Arbitration: A Study of Vacated Arbitration Awards in the New York Appellate Division*, 74 ALB. L. REV. 983 (2010–11). Although only a handful of the cases concerned teacher arbitration, the outer scope of her sample was unclear based on contradictory indicators. Moreover, her vacatur rate is imprecise, because she failed to differentiate inconclusive outcomes and those where the court partially vacated or modified the arbitration award.

[40] Michael H. LeRoy & Peter Feuille, *As the Enterprise Wheel Turns: New Evidence on the Finality of Labor Arbitration Awards*, 18 STANFORD L. & POL'Y REV. 191 (2007) ...1990). However, because their priority was on federal court behavior rather than ultimate outcome rates, they apparently counted the cases that had decisions at both levels separately, thus causing some overlap and double-counting in comparison to limiting the analysis to the final decision.

[41] Michael H. LeRoy & Peter Feuille, *Private Justice in the Shadow of Public Courts: The Autonomy of Workplace Arbitration Systems*, 17 OHIO ST. J. ON DISP. RESOL. 19, 47 (2001); Michael H. LeRoy & Peter Feuille, *The* Steelworkers Trilogy *and Grievance Arbitration Appeals: How the Federal Courts Respond*, 13 INDUS. REL. L.J. 78, 100-01 (1992).

[42] For example, the scope changed with regard to arbitrability cases, and the treatment of partial vacaturs was not entirely consistent.

Neither of these studies specifically analyzed, or even included a significant number of, court decisions reviewing teacher-board grievance arbitration awards. Much earlier, in the only empirical analysis I found specific to this subject, McKinney and Place identified 189 grievance arbitration cases that resulted in judicial appeal in the Westlaw database for the period 1982–1992.[43] They reported that the court confirmed the arbitration award in 66% of the cases and vacated the award entirely or in substantial part in the remaining 34% of the cases.[44] As a result of these decisions, they found, without explanation of the basis for these terms, that the grievant "won" in 55% of the cases, with the remaining 45% in favor of the school district. They also reported that the most frequent issues were dismissal (23%) and duty assignment (22%).

III. PRESENT ANALYSIS

In the light of the absence of recent research concerning court decisions specific to not only vacatur but also, with due differentiation, arbitrability in the context of collective bargaining agreements (CBAs) between school districts and teacher organizations, this section provides a systematic analysis of this case law, with a primary focus on the judicial outcomes. Additionally based on the empirical backdrop of prior research, the design countered the following problematic methodological limitations: (1) employing an overly simplistic or less than objective outcome categorization; (2) including employment arbitration and or other non-CBA cases; (3) double counting within the sample of cases; and (4) limiting the scope to a single issue or jurisdiction.

A. Method

The data collection was based primarily on a Boolean search of the Westlaw database for a ten-year period ending August 1, 2015. The second, supplementary source consisted of the court decisions cited within the relevant parts of the initially selected cases.

---

[43] Joseph R. McKinney & W. William Place, *Judicial Review of Arbitration Awards in the Educational Sector*, 82 EDUC. L. REP. 749 (1993). The description of their data collection and selection procedure was cryptic, referring only to a "topical approach." For example, it is not at all clear as to whether the sample was limited to teacher cases and what the specific exclusions were.

[44] By implication, the upheld category included cases in which the court vacated the award in less than substantial part, although they did not provide any explanation as to the operational meaning of "substantial."

The overall criterion for selection was that the case addressed arbitrability and/or vacatur of grievance arbitration under a teacher-board CBA. Thus, the exclusions included, for example, (1) case dispositions on threshold adjudicative grounds not specific to arbitrability or vacatur, such as lack of jurisdiction or exhaustion; (2) cases concerning nonbinding grievance arbitration; (3) cases arising in the postsecondary education context; (4) cases concerning other school employees; (5) cases limited to separable federal or state claims subsequent to arbitration; and (6) cases limited to costs or attorney's fees for the arbitration.

This systematic search and selection process yielded an ample sample of 110 pertinent final court decisions.[45] A table of the decisions and their coded variables, along with a more detailed version of the literature review, methodology, and results are available in a forthcoming law review article.[46]

B. Findings and Conclusions

The decisions, which were almost entirely in state courts, arose in a total of eighteen states. The jurisdictional frequency distribution revealed that the leading states were as follows in rank order:

1. Pennsylvania - 31 cases
2. New York - 18 cases
3. New Jersey - 10 cases
4. Ohio - 7 cases

5(tie). Massachusetts, Tennessee, and Wisconsin - 5 cases each

The frequency distribution by issue category was, in rank order, as follows:

1. adverse action, especially termination - 51 cases
2. benefits, especially retirement - 17 cases
3. pay, especially salary schedule placement - 15 cases
4. workload and assignments, especially extracurricular activities - 12 cases
5. performance evaluation - 7 cases

[45] Although very likely representative due to the carefully comprehensive coverage, these 110 decisions do not constitute the total population of pertinent decisions. Moreover, to avoid both double counting of and superseded rulings within each case, the coding and analysis was limited to the most recent relevant decision for each case.

[46] Perry A. Zirkel, *Judicial Review of Teacher-School Board Grievance Arbitration: An Empirical Analysis*, __ J.L. & EDUC. __ (forthcoming 2016).

The most notable feature about most of these issue categories was the overlay of state laws specific to public employment generally and, even more, the role, responsibilities, and interrelationship of school boards and teachers specifically.

The table presents the overall distribution of the 110 cases according to judicial outcomes in terms of rulings and, on a net effect basis, in terms of cases.[47] For the sake of uniformity, the rulings for both arbitrability and vacatur are presented in a Yes-No format, necessitating two differentiation adjustments for the vacatur rulings: (1) adding an intermediate outcomes category for the "inconclusive" and "modify" rulings; and (2) using as the frame of reference for the Yes-No outcomes the "confirm," rather than from the "vacate," side of the spectrum. In addition, because some of the cases addressed both arbitrability and vacatur, the total number of the rulings in the first two rows is 120, whereas third row's tabulation of the net effect of these overlapping rulings on a case-by-case basis totals 110.

Table: Distribution of Judicial Outcomes

	Yes	**Inter-mediate**	**No**
Arbitrable? (n=49 rulings)	28 (57%)	--	21 (43%)
Confirm? (n=71 rulings)	53 (75%)	8 (11%)	10 (14%)
Net Effect (n=110 cases)	53 (48%)	26 (24%)	31 (28%)

A review of the Table reveals that arbitrability and vacatur have distinctly different outcomes distributions[48] and that they overlap in their effect. More specifically, the skew in favor of arbitration is less pronounced for the arbitrability (i.e., Arbitrable row) than for the vacatur (i.e., Confirm row) rulings, while the net effect—in contrast with the prior research's focus on vacatur alone—is that only half of the teacher-board arbitration awards survive undisturbed when

[47] Because this analysis, unlike some of the previous studies, included arbitrability and because both arbitrability and vacatur arose in some cases, the net effect on a case-by-case basis shows whether the court upheld or, instead, negated or modified the award.

[48] A major difference is that the skew in favor of arbitration is less pronounced for the arbitrability (i.e., Arbitrable row) than for the vacatur (i.e., Confirm row) rulings.

appealed to court. Conversely, the other half of the cases were almost evenly split between those nullified based on nonarbitrability or vacatur and, to a lesser extent, those in the intermediate outcomes category as a result of being either modified or subject to further proceedings.

Further analysis of the decisions revealed that in reaching these outcomes, the courts applied various approaches or tests to arbitrability and vacatur, respectively. For substantive arbitrability,[49] although the courts in various states used the *Warrior & Gulf* positive assurance test or a similar strong pro-arbitrable presumption, others effectively reached a much more neutral posture by (a) injecting law/public policy into the test, (b) applying the state education code, collective bargaining, or other external law on a more ad hoc basis, or (c) simply subjecting the CBA to straightforward contractual interpretation without any pronounced presumption. For vacatur, the Trilogy-type presumption, whether specifically the essence test or an approximate equivalent, was the prevailing approach. However, alternative or—more frequently—additional approaches prominently included (1) the FAA/UCC standard of exceeding authority per the state's vacatur statute, (2) the contrary to law standard and/or (3) the public policy exception. Moreover, as with arbitrability, the vacatur rulings also conspicuously reflected the direct influence of various and sometimes conflicting laws external to the CBA. In contrast, the various statutory grounds specific to the conduct of the arbitration were notable in their complete absence.

Finally, the outcomes notion of winners and losers requires significant caveats. For arbitrability, the initiator of judicial review was teacher or union in almost a third of the cases, and, in any event, the outcome depended on whether the challenge was based on procedural and/or substantive arbitrability, whether it arose at the pre- or post-arbitration stage, and what the CBA language and the jurisdictional approach were. Moreover, although a court's "no" ruling conclusively defeated the grievant, a "yes" ruling alone amounted to an inconclusive victory for the grievant, depending on further proceedings if issued at the prearbitration stage and the vacatur rulings at the postarbitration stage.

For the vacatur rulings, the caveats are even more tempering in terms of winning and losing. First, contrary to the simple stereotype of the challenge being to a complete arbitral victory for the grievant,

[49] In contrast, procedural arbitrability was rather low frequency and high deference.

the outcomes of the challenged arbitration awards constituted a mixture of denials, partial (i.e., compromise) awards, and full awards. Second, and as interrelated matter, the challenging party in almost one third of the vacatur rulings was the was the union or, in direct opposition to each other, both parties. Third, the mixture of outcomes at the court level, particularly the bottom line effect, of a quarter of the cases in the intermediate limbo category, warns against the stereotype duality of winning or losing.

C. Implications and Recommendations

For frequency, although the predominance of Pennsylvania and New York, together accounting for almost half of the cases, is not surprising in light of their strong unionized tradition and their high overall litigation rates, the non-negligible total number of cases during the most recent ten years is contrary to the prevailing perception that arbitration is final. For outcomes, the results are even more significant in terms of countering the prevailing perception, at least in the teacher-board CBA grievances, that judicial appeal of arbitration is futile. First, arbitrability, whether arising as a prearbitration or postarbitration challenge, was fatal at a 43% rate, perhaps partially attributable to carefully selective challenges but nevertheless contrary to the pronounced pro-arbitral presumption. Second, vacatur, when including modified and inconclusive rulings, had a separable rate of 25%. Finally, their combined net effect left approximately half of the cases as nullified, modified, or in limbo for further arbitral or judicial proceedings.

These results, along with the findings concerning the various approaches or standards for arbitrability and vacatur, highlighted the key distinction for this sector: a multi-level panoply of laws is inextricably influential in a substantial number of the K–12 teacher-board grievance cases. Although the Supreme Court initially opened the door in the commercial context, the historic Meltzer-Howlett debate about the use of "external law" is now moot for this K-12 teacher board collective bargaining context.[50] More specifically, this

[50] Extending the scope to public sector cases in general, a colleague concluded:

> The courts have thrust on labor arbitrators a new role as substitute for litigation in the adjudication of public law claims. Effectively, Howlett has won the debate. Arbitrators can no longer apply the contract and ignore the law.

Martin Malin, *The Evolving Schizophrenic Nature of Labor Arbitration*, 2010 J. DISP. RESOL. 57.

panoply includes not only laws applicable to employers generally (e.g., FMLA and ADEA) and to public sector employers specifically (e.g., First Amendment and public employee CB laws), but also—most distinctively—laws exclusively for K–12 school districts (e.g., education codes, teacher dismissal and certification laws, and teacher-board CB acts).

The primary lesson for practitioners, particularly party representatives and arbitrators in K–12 teacher grievance arbitrations, is to take into special consideration (1) the state-specific standards for both arbitrability and vacatur, (2) the oft-neglected significance of substantive arbitrability at both the pre- and post-arbitration stages, and (3) the distinctive influence of external law, especially in terms of state statutes and judicial precedents specific to the teacher-board relationship.

However, perhaps the most significant message is for legal scholars in terms of the need for more extensive and nuanced empirical research to inform the ongoing debate about judicial review of arbitration awards. More specifically, the recommended research needs to not only more carefully (1) differentiate and compare the types and contexts of arbitration, and (2) the measurement of outcomes including the overlapping effects of arbitrability and vacatur.

IV. POST-SCRIPT

My initial interest in this topic arose earlier in my 40-year career as a part-time grievance arbitrator in various collective bargaining contexts. In one of the very few cases where one or both parties sought judicial review, the court overturned my award, which was a carefully crafted compromise in a teacher termination case under a collective bargaining agreement's rather typical just cause clause.[51] The school board charged the teacher with writing love letters to two female high school students. The parties agreed on two issues, including the corollary concerning the remedy.[52] I found that the

[51] Manheim Cent. Educ. Ass'n v. Manheim Cent. Sch. Dist., 572 A.2d 31 (Pa. Commw. Ct. 1990).

[52] *Id*. at 34 n.4:

1. Whether the grievant was discharged for just cause? If not, what shall the remedy be?

2. Whether the grievant was improperly denied accumulated sick leave and a medical sabbatical? If so, what shall the remedy be?

teacher had written the letters, although he did not intimate, much less commit, any physical contact with them.[53] In light of expert testimony showing the psychological effects of his exceptionally stressed situation[54] and various mitigating circumstances that arbitrators often apply in just cause cases,[55] I ordered a lengthy unpaid suspension, the opportunity to use his accumulated sick leave for psychotherapy, and an offer of resignation. However, in a 2-to-1 ruling the court effectively concluded that once I found that the grievant-teacher had written the love letters, which constituted immorality under Pennsylvania's school code,[56] just cause was established and I was without authority to order a remedy. The dissent concluded that my award met the applicable approach—the essence test. However, whether viewed as the exceeds authority, the public policy, and/or the contrary to law standard, the majority opinion in this case was a harbinger of this empirical analysis, which provides a more comprehensive, current, and complementary perspective to traditional legal scholarship.

[53] Although teachers' abuse of their special power position with students is of historic concern, this case arose before the relatively recent cognizance of sexual harassment in public schools. See, e.g., Gebser v. Lago Vista Indep. Sch. Dist., 524 U.S. 274 (1998); Franklin v. Gwinnett Cnty. Pub. Sch., 503 U.S. 60 (1992).

[54] *Id*. at 33:

> His father died, he became estranged from his brother due to his mother's debilitating illness, his wife had an affair and asked for a divorce, he placed his mother in a nursing home, he moved out of his house at his wife's insistence, and his son went off to college, and his mother died of cancer.

[55] *Id*. at 33 (identifying his long and exemplary record and esteemed reputation among both students and staff members).

[56] Pennsylvania uses the traditional community-based exemplar, or teacher on a pedestal, rather than the modern nexus, or actual effect on teaching, model for immorality. See, e.g., Bonatesta v. N. Cambria Sch. Dist., 48 A.3d 552, 558 (Pa. Commw. Ct. 2012) (citing Horosko v. Sch. Dist. of Mt. Pleasant Twp., 6 A.2d 866, 868 (Pa. 1939)).