eJEP
eJournal of Education Policy

The Education Value-Added Assessment System (EVAAS) on Trial: A Precedent-Setting Lawsuit with Implications for Policy and Practice

Audrey Amrein-Beardsley, PhD
Professor of Educational Policy and Evaluation
Arizona State University
Address: PO Box 871811, Tempe, AZ 85287-1811
602-561-4731
audrey.beardsley@asu.edu

Spring 2019

Abstract
In *Houston Federation of Teachers (Plaintiff) v. Houston Independent School District (Defendant)*, plaintiffs contested how teachers were being evaluated using a value-added model (VAM) as part of the districts' teacher accountability system. The VAM at issue was the Education Value-Added Assessment System (EVAAS), with potential consequences at issue involving teacher termination as per "insufficient student academic growth" reflected by teachers' EVAAS estimates. In this piece I describe this case, as well as a recent ruling of significance regarding teachers' Fourteenth Amendment rights as pertinent to this case, but also with broad implications for others using the EVAAS (and perhaps other VAMs) for similar teacher accountability purposes.

Keywords: Accountability, Educational Policy, Educational Reform, Law/Legal, Teacher Assessment

Introduction

As per *Education Week* (2015), there are approximately 15 lawsuits in the U.S. in which plaintiffs are contesting how teachers are being evaluated using value-added models (VAMs) as the primary component of states'/districts' contemporary teacher evaluation systems (U.S. Department of Education, 2014). Of these lawsuits, perhaps the most controversial and highest profile was in Texas: *Houston Federation of Teachers (Plaintiff)* v. *Houston Independent School District (Defendant)*, U.S. District Court, Southern District of Texas, Houston Division.

The Houston Independent School District (HISD) is the largest district in Texas and the seventh largest district in the country. The majority of the students in the district are from high-needs backgrounds, with 76% qualifying for federal free or reduced lunches, 64% of students labeled at risk, and 60% classified as English Language Learners (ELLs), Limited English Proficiency

(LEP), or bilingual (HISD, 2016). HISD also continues to be one of the lowest performing, or rather "worst" districts in the state and nation (Leicht, 2014).

Consequently, in 2007, HISD contracted (at $500,000 to $680,000 per year) with SAS Institute Inc. to use its Education Value-Added Assessment System (EVAAS) – "the most comprehensive reporting package of value-added metrics available in the educational market" (SAS Institute Inc., 2017) – for district reform purposes. The goal was to better hold district teachers accountable for the academic progress their students made (or did not make) from year-to-year to ultimately help alter the district's history of poor performance.

Correspondingly, by 2010, no other district or state using the EVAAS, or any other VAM for that matter, was using a VAM for more consequential decision-making purposes than HISD (Corcoran, 2010; Harris, 2011; Otterman, 2010). In 2011, for example, HISD did not renew 221 of its teachers' contracts given they demonstrated "a significant lack of student progress attributable to the educator" or "insufficient student academic growth reflected by [their EVAAS] value-added scores" (HISD, 2011). This, along with other alleged (mis)uses of the EVAAS (e.g., using EVAAS estimates for 50%-100% of teachers' overall evaluation scores, with other contractual, termination, and merit pay decisions also attached to EVAAS estimates), brought forward this lawsuit. Put differently, the highly consequential nature of Houston's teacher evaluation system as allegedly, largely, and sometimes solely based on the EVAAS may be literally what landed this district in a U.S. District Court.

EVAAS

The EVAAS was developed in Tennessee in the 1980s by the late Dr. William L. Sanders as the Tennessee Value-Added Assessment System (TVAAS). Now more generally known as the EVAAS, the system links student tests scores to their teachers to more objectively make causal determinations about teachers' impacts on student achievement over time. Students' levels of growth are measured by changes in students' large-scale standardized achievement test scores over time, while controlling for students' test histories (and other characteristics upon state or district request). Included are all eligible students with preferably at least three and up to five years of prior test scores across multiple subject areas, grade levels, and years (even if incomplete or fragmented). Growth scores are then aggregated to yield teacher-level value-added estimates. Depending on where teachers' value-added estimates fall, as compared to other similar teachers to whom they are compared, teacher-level value-added determinations are then made. Teachers whose students grow significantly more than the average teacher, and who surpass their projected levels of growth are identified as adding value. Teachers whose students grow significantly less than the average teacher, and who fall short of projected levels of growth are identified as not adding value. Teachers whose students grow at rates that are not statistically different from average (i.e., falling within one standard deviation of the mean) are classified as "not detectibly different" (NDD).

As such, and as marketed, the EVAAS offers states, districts, and schools "precise, reliable and unbiased results that go far beyond what other simplistic [value-added] models found in the market today can provide" (SAS Institute Inc., 2017). Whether this is realized in practice, however, is uncertain and of note in this case which has since set precedent for others.

HFT v. HISD

With the filing of *Houston Federation of Teachers (Plaintiff)* v. *Houston Independent School District (Defendant)*, plaintiffs alleged that EVAAS output are unreliable and imprecise, as contrary to the above marketing claim (see also Chiang, McCullough, Lipscomb, & Gill, 2016; Martinez, Schweig, & Goldschmidt, 2016; Yeh, 2013). Plaintiffs alleged that the EVAAS is unfair in that teachers of some students, grades, or subject areas are more susceptible to the consequences attached to EVAAS estimates versus others (see also Ballou & Springer, 2015; Baker, Oluwole, & Green, 2013; Harris, 2011). Plaintiffs alleged that teachers are being evaluated via the EVAAS using tests that do not match the subject areas or curricula they teach (see also American Educational Research Association (AERA) Council, 2015; Collins, 2014; Eckert & Dabrowski, 2010). And plaintiffs alleged that the EVAAS system fails to control for biasing factors that impact how well teachers perform but that are outside of teachers' control (e.g., race, socioeconomic, and language factors), as also contrary to the above marketing claim (see also Koedel, Mihaly, & Rockoff, 2015; Newton, Darling-Hammond, Haertel, & Thomas, 2010; Rothstein, 2009, 2010).

Plaintiffs also alleged that the EVAAS is opaque, incomprehensible (e.g., a "black box" model) and, hence, difficult to use to improve upon teachers' professional practices (i.e., EVAAS estimates are not "actionable;" see also Eckert & Dabrowski, 2010; Gabriel & Lester, 2013; Goldring et al., 2015). Consequently, plaintiffs also alleged that teachers' Fourteenth Amendment due process rights were being violated because teachers do not have adequate opportunities to change their professional practices as a result of accessing, understanding, and using their EVAAS reports. Likewise, HISD teachers whose contracts were to be terminated on the basis of EVAAS scores could neither independently verify nor challenge the scores or their resulting contract terminations (see also Ballou & Springer, 2015).

The main issue here is that most VAM-based estimates do not seem to make sense to those at the receiving ends, whereas teachers do not seem to understand the models being used to evaluate them; hence, teachers are reportedly unlikely to use the output for formative purposes. More expressly, value-added data reports are often described by practitioners as confusing, not comprehensive in terms of the key concepts and objectives taught, ambiguous in terms of teachers' efforts at both the student and aggregate or composite levels, and often received months after students leave teachers' classrooms (Eckert & Dabrowski, 2010; Gabriel & Lester, 2013; Goldring et al., 2015).

More generally, HISD teachers expressed similar concerns noting that they are learning little about what they did (in)effectively to yield particular EVAAS estimates and how they might use their EVAAS estimates to improve their instruction as a result (Collins, 2014). Teachers in North Carolina reported that they are "weakly to moderately" familiar with their EVAAS data and how to use them (Kappler Hewitt, 2015). Eckert and Dabrowski (2010) also demonstrated that in Tennessee that teachers did not understand use their TVAAS (i.e., the Tennessee version of the EVAAS, noted prior) data in order to use them to improve upon their professional practice (see also Harris, 2011). These assertions were made despite systems also put into place to support data comprehension and use (e.g., as supported by an EVAAS partner – Battelle for Kids).

Notwithstanding, this is problematic in that another of the primary strengths of the EVAAS, as also advertised, is its wealth of "valuable diagnostic information" (SAS Institute Inc., 2017) accumulated for such formative uses and purposes (see also Sanders, Wright, Rivers, & Leandro, 2009). Although at the same time, and in disagreement, EVAAS modelers sometimes make "no apologies for the fact that [their] methods [are] too complex for most of the teachers whose jobs depended on them to understand" (Carey, 2017; see also Gabriel & Lester, 2013).

The unfortunate truth remains that EVAAS estimates may tell teachers little about how they might understand and then use their own data so as to improve upon their instructional practice, to ultimately improve student learning and achievement over time. This turned out to be the most critical claim asserted by plaintiffs in this case.

Court Ruling on Protective Order

First, of note in this regard was that despite Public Information Act requests and a litigation subpoena submitted by the plaintiffs for this case, EVAAS proprietors provided very limited access to the data used to evaluate HISD teachers in response. HISD (involving SAS Institute Inc.) provided general descriptions of their statistical procedures and formulas, technical reports, and the like, but they did not release the decision rules and secured algorithms, source codes, etc. needed to actually examine and assess the accuracy of HISD teachers' EVAAS estimates.

After much negotiation, though, SAS Institute Inc. agreed to give one of two of the plaintiffs' expert witnesses access to this protected information, although this was to be done using a highly controlled and monitored procedure and place (i.e., the expert witness was permitted to view the EVAAS program code on a laptop computer in one of the SAS Institute Inc.'s lawyer's offices). This scenario was later demarcated by this case's judge as "extremely restrictive access" (U.S. District Court, 2016, p. 5). Nonetheless, the expert witness found that even with the access (s)he was provided, (s)he was not able to reproduce the EVAAS calculations so that they could be verified for HISD teachers.

After this expert witness wrote his/her findings into his/her affidavit, and both plaintiffs' expert witnesses' key findings from their affidavits went public via an American Federation of Teachers

(AFT)/Houston Federation of Teachers (HFT) press release and subsequent blog post (see VAMboozled, 2016), this eventually led to SAS Institute Inc. claiming an alleged violation of the protective order in place to protect their proprietary system throughout this case. SAS Institute Inc. argued, more specifically, that what this expert witness wrote into his/her affidavit, and its public dissemination, exposed information covered by the protective order associated with the EVAAS. In the end, though, the judge ruled that SAS Institute Inc. interpreted their protective order "too broadly" (p. 4), adding that the "overly broad interpretation urged by [SAS Institute Inc.] would inhibit legitimate discussion about the lawsuit" (p. 6); hence, the sanctions that SAS sought against the plaintiffs were denied.

This incident is relevant here, with specific regards to transparency and data use, also in that it foreshadowed and played itself out in the judge's later ruling on whether there was sufficient evidence to allow the plaintiffs to proceed to trial on their due process claims premised on such secrecy. Hence, what was ultimately questioned was whether teachers, without having near the access that the aforementioned expert witness had, were able to understand and use, or replicate or challenge their scores, themselves, so as to use them to improve upon their professional practice or defend themselves, respectively, if they felt that their scores were produced incorrectly, were based on faulty data, were biased by the types of students they taught, etc.

Court Ruling on Teachers' Fourteenth Amendment Rights

One year later in May of 2017, the defendant moved for summary judgment on all matters pertaining to the case (noted prior). After the judge examined all evidence presented by each side, the judge ruled on all but one of the claims in the defendant's favor (e.g., that evaluating teachers using their value-added scores was "unconstitutionally unreasonable" was unsuccessful, as it was in a similar case pertaining to EVAAS use in Tennessee).

However, the judge notably ruled that the plaintiffs did have sufficient evidence to proceed to trial on their claims that the use of EVAAS in Houston to terminate teacher contracts was a violation of their Fourteenth Amendment due process protections (i.e., no state or in this case district shall deprive any person of life, liberty, or property, without due process).

As taken from the judge's summary judgment opinion (U.S. District Court, 2017), the plaintiffs challenged the use of the EVAAS for such teacher accountability purposes under their Fourteenth Amendment right to "procedural due process, due to lack of sufficient information to meaningfully challenge terminations based on low EVAAS scores" (p. 4), given "due process is designed to foster government decision-making that is both fair and accurate" (p. 10). That is, the judge ruled that "HISD's value-added appraisal system pose[d] a realistic threat to deprive plaintiffs of constitutionally protected property interests in employment" (p. 9); hence, this case was to move forward to trial.

The judge added that because HISD does not calculate teachers' EVAAS estimates, but rather delegates these computational tasks to the same third-party vendor, SAS Institute Inc., this is also problematic. HISD conceded to this being problematic. Further, HISD acknowledged that "any effort by teachers to replicate their own scores, with the limited information available to them, [would] necessarily fail" (U.S. District Court, 2017, p. 13). This was confirmed by the aforementioned plaintiffs' expert witness who, again, was "unable to replicate the scores despite being given far greater access to the underlying computer codes than is available to an individual teacher" (U.S. District Court, 2017, p. 13).

Likewise, because teachers' EVAAS estimates "might be erroneously calculated for any number of reasons, ranging from data-entry mistakes to glitches in the computer code itself" (U.S. District Court, 2017, p. 13), and given there are no verification or auditing processes in place at the district-level, this contributed to the court's noted concerns. After all, "[a]lgorithms are human creations, and subject to error like any other human endeavor." That HISD also acknowledged that once mistakes occur, "it will not [and perhaps cannot] be promptly corrected" (U.S. District Court, 2017, p. 13), this was underscored in the judge's registered concerns.

While Ballou & Springer (2015) also critiqued EVAAS's practices by which EVAAS modelers revise teachers' aggregate scores to get at more precision as more information becomes available about teachers' students over time, which also causes individual teachers' EVAAS scores to change retroactively, also of explicit issue here was that while this occurs at a system-wide level, individual teacher-level estimates will also not be recalculated otherwise.

> Once EVAAS analyses are completed, any re-analysis can only occur at the system level. What this means is that if we change information for one teacher, we would have to re-run the analysis for the entire district, which has two effects: one, this would be very costly for the district, as the analysis itself would have to be paid for again; and two, this re-analysis has the potential to change all other teachers' reports. (HISD, 2015, p. 6)

In response to this excerpt, the judge wrote:

> The remarkable thing about this passage is not simply that cost considerations trump accuracy in teacher evaluations, troubling as that might be. Of greater concern is the house-of-cards fragility of the EVAAS system, where the wrong score of a single teacher could alter the scores of every other teacher in the district [see also Ballou & Springer, 2015]. This interconnectivity means that the accuracy of one score hinges upon the accuracy of all. Thus, without access to data supporting all teacher scores, any teacher facing discharge for a low value-added score will necessarily be unable to verify that her [sic] own score is error-free. (U.S. District Court, 2017, p. 14)

Hence, and "[a]ccording to the unrebutted testimony of [the] plaintiffs' expert [witness], without access to SAS's proprietary information – the value-added equations, computer source codes, decision rules, and assumptions – EVAAS scores will remain a mysterious 'black box,' impervious to challenge" (U.S. District Court, 2017, p. 17).

While HISD defended itself noting that it should not have to replicate EVAAS scores "down to the last decimal point" (U.S. District Court, 2017, p. 17), the judge responded that because EVAAS scores are calculated to the second decimal place, "an error as small as one hundredth of a point could spell the difference between a positive or negative EVAAS effectiveness rating, with serious consequences for the affected teacher" (U.S. District Court, 2017, p. 18). Consequently, the judge concluded that HISD teachers "have no meaningful way to ensure correct calculation of their EVAAS scores, and as a result are unfairly subject to mistaken deprivation of constitutionally protected property interests in their jobs" (U.S. District Court, 2017, p. 18).

Most recently (i.e., October of 2017), the parties reached a settlement agreement. As part of the agreement, HISD agreed to refrain from using VAMs, including the EVAAS, to terminate teachers' contracts as long as the VAM score is "unverifiable." More specifically, as written into the federal suit settlement:

> HISD agrees it will not in the future use value-added scores, including but not limited to EVAAS scores, as a basis to terminate the employment of a term or probationary contract teacher during the term of that teacher's contract, or to terminate a continuing contract teacher at any time, so long as the value-added score assigned to the teacher remains unverifiable. (State of Texas, 2017, p. 2; see also AFT, 2017)

HISD also agreed to create an "instructional consultation subcommittee" to more inclusively and democratically inform HISD's teacher appraisal systems and processes, and HISD agreed to pay the Texas AFT $237,000 in its attorney and other legal fees and expenses (State of Texas, 2017, p. 2; see also AFT, 2017).

Conclusions and Implications

On this allegation alone, the efficacious use of the EVAAS in HISD, perhaps elsewhere and for other VAMs elsewhere, as this ruling is an unprecedented development in VAM litigation, is now (or should now be) of more interest and concern. As per the judge, accordingly, "[w]hen a public agency adopts a policy of making high stakes employment decisions based on secret algorithms incompatible with minimum due process, the proper remedy [may be] to overturn the policy" (U.S. District Court, 2017, p. 18).

Likewise, because this is the first case to yield such a decision, it is important for all in the academy of education to note as, again this also has broad implications for educational policy and practice. For the immediate future, states and districts that have adopted the EVAAS for similar teacher evaluation purposes might ensure that their teachers' due process rights are also not in jeopardy. Because the EVAAS system is proprietary, and reliant upon contracts similar to the now former contract in place in Houston (i.e., in 2016 the HISD school board voted to suspend its contract with SAS Institute Inc., although this did not resolve this case), however, other states and districts may be in similar situations given SAS Institute Inc.'s standard contractual policies and procedures. Given the likelihood of this being the case, states and districts might otherwise consider not attaching consequential decisions to teachers' EVAAS estimates, given that the severity of the consequences attached may concurrently heighten the importance of teachers' due process rights in these regards. On that note, whether states and districts continue to use the EVAAS for low-stakes purposes might also be reconsidered, given the implications of the court's standing concerns about the extent to which, even for low-stakes purposes, teachers might understand their EVAAS estimates so as to use them to improve upon their practice.

States and districts using other VAMs might also consider the case of the EVAAS in Houston to assess not only how accessible and comprehensible their VAM-based data are, but also how likely teachers might be to also use their data to defend themselves, especially against specious consequences, and also to improve upon their professional practice. This is also important as many states and districts continue to employ similar teacher evaluation systems, still as based in significant part on VAMs, despite the fact that as per the Every Student Succeeds Act (ESSA) states are no longer required (or incentivized) to engage is these policies (see also Kraft & Gilmour, 2017). It is also important not only that others understand what is (still) happening in this area of research (e.g., how educational research might have a direct and measurable impact on policy and practice writ large; Furlough, 2010; Green, 2000; Shulenburger, 2005), but how this is happening in real time via America's judicial system, and impacting educational policy and practice as a result.

References

American Educational Research Association (AERA) Council. (2015). AERA statement on use of value-added models (VAM) for the evaluation of educators and educator preparation programs. *Educational Researcher, 44*(8), 448-452. doi:10.3102/0013189X15618385 Retrieved from http://edr.sagepub.com/content/early/2015/11/10/0013189X15618385.full.pdf+html

American Federation of Teachers (AFT). (2017). Federal suit settlement: End of value-added measures in Houston. *Education News*. Retrieved from https://www.educationviews.org/federal-suit-settlement-value-added-measures-houston/

Baker, B. D., Oluwole, J. O., & Green, P. C. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the Race-to-the-

Top era. *Education Policy Analysis Archives, 21*(5), 1-71. Retrieved from http://epaa.asu.edu/ojs/article/view/1298

Ballou, D., & Springer, M. G. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher, 44*(2), 77-86. doi:10.3102/0013189X15574904

Carey, K. (2017). The little-known statistician who taught us to measure teachers. *The New York Times*. Retrieved from https://www.nytimes.com/2017/05/19/upshot/the-little-known-statistician-who-transformed-education.html?_r=0

Chetty, R., Friedman, J., & Rockoff, J. (2014). Measuring the impact of teachers II: Evaluating bias in teacher value-added estimates. *American Economic Review, 104*(9), 2593-2632. doi:10.3386/w19424

Chiang, H., McCullough, M., Lipscomb, S., & Gill, B. (2016). Can student test scores provide useful measures of school principals' performance? Washington DC: U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/pubs/2016002/pdf/2016002.pdf

Collins, C. (2014). Houston, we have a problem: Teachers find no value in the SAS Education Value-Added Assessment System (EVAAS®). *Education Policy Analysis Archives.* Retrieved from http://epaa.asu.edu/ojs/article/view/1594

Corcoran, S. P. (2010). *Can teachers be evaluated by their students' test scores? Should they be? The use of value-added measures of teacher effectiveness in policy and practice*. Providence, RI: Annenberg Institute for School Reform. Retrieved from http://annenberginstitute.org/publication/can-teachers-be-evaluated-their-students%E2%80%99-test-scores-should-they-be-use-value-added-mea

Eckert, J. M., & Dabrowski, J. (2010, May). Should value-added measures be used for performance pay? *Phi Delta Kappan, 91*(8), 88-92.

Education Week. (2015). *Teacher evaluation heads to the courts*. Retrieved from http://www.edweek.org/ew/section/multimedia/teacher-evaluation-heads-to-the-courts.html

Furlough, M. (2010). Open access, education research, and discovery. *Teachers College Record, 112*(10), 2623-2648. Retrieved from http://www.tcrecord.org/Content.asp?ContentId=15874

Gabriel, R. & Lester, J. N. (2013). Sentinels guarding the grail: Value-added measurement and the quest for education reform. *Education Policy Analysis Archives, 21*(9), 1-30. Retrieved from http://epaa.asu.edu/ojs/article/view/1165

Goldring, E., Grissom, J. A., Rubin, M., Neumerski, C. M., Cannata, M., Drake, T., & Schuermann, P. (2015). Make room value-added: Principals' human capital decisions and the emergence of teacher observation data. *Educational Researcher, 44*(2), 96-104. doi:10.3102/0013189X15575031

Green, M. F. (2000). Bridging the gap: Multiple players, multiple approaches. *New Directions for Higher Education, 110,* 107-113.

Harris, D. N. (2011). *Value-added measures in education: What every educator needs to know*. Cambridge, MA: Harvard Education Press.

Houston Independent School District (HISD). (2011). *Board of education workshop 2011-12 budget update* [PowerPoint slides]. Houston, TX.

Houston Independent School District (HISD). (2015). *EVAAS®/value-added frequently asked questions.* Houston, TX. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwiwjaa25JjWAhVI72MKHTKjD2gQFggmMAA&url=http%3A%2F%2Fstatic.battellefo

rkids.org%2Fdocuments%2FHISD%2FEVAAS-Value-Added-FAQs-Final-2015-02-02.pdf&usg=AFQjCNGwMwoxzAgLhv0KOehOV7-8PJPE7g
Houston Independent School District (HISD). (2016). 2015-2016 Facts and figures. Houston, TX. Retrieved from http://www.houstonisd.org/achievements
Kappler Hewitt, K. (2015). Educator evaluation policy that incorporates EVAAS value-added measures: Undermined intentions and exacerbated inequities. *Education Policy Analysis Archives, 23*(76), 1-49. Retrieved from http://epaa.asu.edu/ojs/article/view/1968
Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review, 47*, 180–195. doi: 10.1016/j.econedurev.2015.01.006
Kraft, M. A, & Gilmour, A. F. (2017). Revisiting the Widget Effect: Teacher evaluation reforms and the distribution of teacher effectiveness. *Educational Researcher, 46*(5) 234-249. doi:10.3102/0013189X17718797
Leicht, A. (2014). Plenty of Houston-area schools are among Texas's worst, according to TEA list. *Houston Press*. Retrieved from http://www.houstonpress.com/news/plenty-of-houston-area-schools-are-among-texass-worst-according-to-tea-list-6719742
Martinez, J. F., Schweig, J., & Goldschmidt, P. (2016). Approaches for combining multiple measures of teacher performance: Reliability, validity, and implications for evaluation policy. *Educational Evaluation and Policy Analysis, 38*(4), 738–756. doi: 10.3102/0162373716666166 Retrieved from http://journals.sagepub.com/doi/pdf/10.3102/0162373716666166
Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010) Value-Added Modeling of Teacher Effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives, 18* (23), 1-27. Retrieved from http://epaa.asu.edu/ojs/article/view/810
Otterman, S. (2010, December 26). Hurdles emerge in rising effort to rate teachers. *New York Times*. Retrieved from http://www.nytimes.com/2010/12/27/nyregion/27teachers.html
Rothstein, J. (2009). Student sorting and bias in value-added estimation: Selection on observables and unobservables. *Education Finance and Policy*, (*4*)*4*, 537-571. doi:10.3386/w14666
Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics*, 175-214. doi:10.1162/qjec.2010.125.1.175
Sanders, W. L., Wright, S. P., Rivers, J. C., & Leandro, J. G. (2009, November). *A response to criticisms of SAS EVAAS*. Cary, NC: SAS Institute Inc. Retrieved from http://www.sas.com/resources/asset/Response_to_Criticisms_of_SAS_EVAAS_11-13-09.pdf
SAS Institute Inc. (2017). *SAS® EVAAS® for K-12: Assess and predict student performance with precision and reliability.* Cary, NC. Retrieved from http://www.sas.com/govedu/edu/k12/evaas/index.html
Shulenburger, D. E. (2005). Public goods and open access. *New Review of Information Networking, 11*(1), 3-11. doi: 10.1080/13614570500268282
State of Texas. (2017). Settlement and full and final release agreement. Houston, TX. Retrieved from http://www.aft.org/sites/default/files/settlementagreement_houston_100717.pdf
U.S. Department of Education. (2014). *States granted waivers from No Child Left*

Behind allowed to reapply for renewal for 2014 and 2015 school years. Washington D.C. Retrieved from http://www.ed.gov/news/press-releases/states-granted-waivers-no-child-left-behind-allowed-reapply-renewal-2014-and-2015-school-years

U.S. District Court. (2016). *Court order.* Stephen Wm Smith, U.S. Magistrate Judge. Retrieved from http://vamboozled.com/wp-content/uploads/2016/03/Court-Order.pdf

U.S. District Court. (2017). *Amended summary judgment opinion.* Stephen Wm Smith, U.S. Magistrate Judge. Retrieved from https://casetext.com/case/houston-fedn-of-teachers-v-houston-indep-sch-dist

VAMboozled. (2016). *Houston lawsuit update, with summary of expert witnesses' findings about the EVAAS.* Tempe, AZ. Retrieved from http://vamboozled.com/houston-lawsuit-update-with-summary-of-expert-witnesses-key-findings-about-the-evaas-vam/

Yeh, S. S. (2013). A re-analysis of the effects of teacher replacement using value-added modeling. *Teachers College Record, 115*(12), 1-35. Retrieved from http://www.tcrecord.org/Content.asp?ContentID=16934