

1. Ainsworth, L. (2003). *Power standards: Identifying the standards that matter the most*. Denver, CO: Advanced Learning Press.

While the concept of value-added assessment does not mention the terminology of “power standards” the two are connected. This book serves as a practical professional development guide that assists practitioners with professional development related to the investigation of state level academic content standards and grade level learning targets. The book articulates a methodology whereby professional development planners can prioritize the most important learning targets for a particular grade level. These learning targets are said to be the most important because they endure, leverage and build a foundation and provide readiness for future learning. The author points out that while all learning standards might be important, those that meet a specific criteria are the most important. The Value Added Assessment System (VAAS) purports to measure student acquisition of grade level learning targets and tracks whether or not individuals are making expected year to year progress on these targets. A professional development sequence that could help provide teachers with insight into teaching the most important learning targets that leverage, endure and provide readiness might very well help them add more value to student achievement.

Relevancy to main topic: Relates to the professional development that might be provided to teachers so that they can become more adept at standards-based education.

2. Ainsworth, L. (2003). *Unwrapping the standards: A simple process to make standards manageable*. Denver, CO: Advanced Learning Press.

This book serves as another practical guide that can assist district level professional developers. In order to help teachers become more familiar with the learning targets they are teaching students, this book helps educators set up professional develop that incorporates a methodology that takes the broad learning targets and breaks them down into their component parts. Learning concepts are divided into Big Ideas and Essential Questions. The Big Ideas allow teaches to focus student lessons and the Essential Questions provide students with advanced organizers regarding what they should expect to know and be able to do at the end of learning. A professional development sequence that could help provide teachers with skills to make the student learning targets clearer might very well help them add more value to student achievement.

Relevancy to main topic: Relates to the professional development that might be provided to teachers so that they can become more adept at standards-based education.

3. Andrejko, L. (2004). Value-added assessment: A view from a practitioner. *Journal of Educational and Behavioral Statistics*, 29 (1), 7-9.

This is a descriptive article that reveals the pros and cons of a value-added assessment system. The author is a superintendent of schools who points out the advantages and disadvantages of a VAAS and calls for more qualitative measures to be included with the VAAS to evaluate teachers. Student test scores should be used for progress measurement first and perhaps teacher evaluation at a later time.

Relevancy to main topic: Serves as a practitioner's perspective on the practical implications of VAAS.

4. Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29 (1), 37-65.

This article appeared in the JEBS issue that was devoted to the topic of value added. It was written to address criticisms from the statistical community; particularly from Kupermintz's assertions in 2003. The Tennessee Value Added Assessment System (TVAAS) model developed by William Sanders is somewhat different compared to other models. In TVAAS, specific controls are not put into place to account for student SES, demographic and other variables that have been found to have an impact on student achievement. Sanders and his colleagues have argued that they have found that controlling for such variables at the student level does not show much difference when one is trying to estimate the impact that the teacher has on student achievement gain; thus the criticism from within the measurement community.

TVAAS is different when compared to other simple fixed-effects estimators. TVAAS was designed to be able to account for missing data in that the student's entire test score history is included and used to estimate expected gains from observed gains; a simple fixed-effects estimator would not do this and as a result might very well need to control for SES and demographics. TVAAS also has the student serve as his or her own baseline. As a result, if a student's previous scores are compared to a current achievement performance, the argument would be that things like SES and demographics

are already accounted for because they existed during the obtainment of each gathering of achievement data.

The authors conclude by indicating that they believe they have addressed the criticisms that TVAAS has received. One unresolved question relates to the composition of the kind of class that a student might be in. TVAAS assumes that SES and demographic “mixing” has taken place when students are in classes and the observations that have taken place using the TVAAS system has been in school systems where mixing has occurred. The authors caution using TVAAS or a form of it in order to measure teacher effects on student achievement gains when insufficient mixing has taken place. If a district has a great deal of social or racial stratification and then seeks to make teacher gain comparisons using drastically different populations, problems might occur. The authors postulate that in such an environment “the school influences achievement through peer effects.”

Relevancy to main topic: This article addresses the criticism of TVAAS’s model head on. While technical and somewhat complicated, the authors do a very good job of summarizing the important conclusions at the end. This article can be used to refute arguments that the TVAAS does not account for SES and demographic influences when making student gain predictions.

5. Black, P. & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, October, 139-48.

This article was perhaps the most important one released on student assessment within the past decade. The article brought to forefront the ultimate purpose of student assessment—it is for learning. It sought out to answer three questions. First, does formative assessment help to raise standards? By reviewing 250 articles by researchers

from around the world, the authors presented very strong support via their meta analysis that simply engaging in formative assessment (a method that entails providing feedback to students) has a positive effect on student achievement. The second question in the article asked whether or not there was room for improvement in the practice of formative assessment. The answer to this question yielded three responses. First, current assessment practices by teachers typically did not promote learning. Second, the grading of student work promoted competition rather than mastery. Finally, feedback from teachers was negative rather than positive or formative. The final third question sought out to answer the question of whether or not there was evidence as to how to improve current formative assessment practices. They concluded that more practical examples of effective formative assessment practices were needed.

Relevancy to main topic: Provides theoretical assumptions backed by data to demonstrate that monitoring the progress of students and providing them with valuable assessment experiences increases learning. Has application in that while a system to measure the value that has been added to student achievement is needed, teaching teachers how to look at data and use it is the next step.

6. Black, P., Harrison, C., Lee, C., Marshall & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom.

This article is an extension and follow up to the 1998 article written by Black and Wiliam. The purpose of it was to provide more information to practitioners about the final question from the 1998 article. That question was, what evidence exists that will help improve student achievement through formative assessment. This article is a review an action research project conducted within two school districts in the country of England. Twenty-four teachers participated in the project and the data obtained that led

to recommendations was gathered by teacher observations, records of meetings, and via interviews with teachers and students. The recommendations that would lead to increased student achievement include: 1) changing the method whereby teachers question students within the classroom so that it includes more wait time and reflects deeper questions; 2) change grading practices so that they rely more on comments and constructive feedback and less on a score; 3) allow time for students to self-assess and peer assess; and 4) encourage students to engage in reflective review of their work (particularly when summative assessments are considered).

Relevancy to main topic: The creation of a VAAS alone does not insure teachers creating quality assessments that help students. This article provides recommendations as to how teachers can engage in best practice by way of assessment creation and feedback.

7. Boyle, B., While, D. & Boyle, T. (2004). A longitudinal study of teacher change: What makes professional development effective? *The Curriculum Journal*, 15 (1), 45-68.

This study was conducted by a group of researchers from England. The researchers review the problems with professional development in that it has often been characterized by teachers as lacking focus and continued direction. They draw similarities between that which has been found in England and the United States. This study sought out to do four things. 1) Investigate the different professional development (PD) models; 2) investigate whether or not continued professional development has strength in changing teacher practices when compared to a PD that is not on-going; 3) is there a relationship between PD and changes in teaching practice; and 4) do certain characteristics of PD have an impact on student achievement. Data were obtained from

approximately 60 schools in England via questionnaire surveys. The authors conclude that at this point they have preliminary findings and qualify it by indicating that this study is longitudinal and data collection will be on-going. Preliminary findings: teachers participate in PD; most PD is long term in nature; based on teacher report, the most powerful PD included sharing teaching practices and observing colleagues; the majority of teachers reported changes in their teaching practices based on PD activities. Further study will be done to look for the strength of relationship between PD and student performance.

Relevancy to main topic: The roll out of an EVAAS will create the need to develop teachers so that they are better able to be good consumers of data and know how to use it to guide their instruction so that it ultimately leads to gains in student achievement. This study provides evidence that these goals are not just evident in the U.S., but elsewhere.

8. Braun, H. I. (2005, September). Using student progress to evaluate teachers: A primer on value-added models. Retrieved June 21, 2006, from <http://www.ets.org/Media/Research/pdf/PICVAM.pdf>

This report serves as a layperson's guide to an overview of the different models associated with value added assessment. Seven questions are raised and include: Why is there such interest in value-added modeling; What is the fundamental concern about VAMSs; What are some specific concerns about treating estimated teacher effects as measures of teacher effectiveness; What value-added models are now in use; How does EVAAS work; What are some of the issues in using student achievement data in teacher evaluation and Where do we stand?

Relevancy to main topic: Technical articles that explain the purpose of a VAAS are important; practitioners guides that explain the reality of it are

equally important. In order for an evaluation model to work effectively, it is important that those who are being evaluated understand how it works.

9. Callender, J. (2004). Value-added student assessment. *Journal of Educational and Behavioral Statistics*, 29 (1), 5.

This is one of the first articles written in a JEBS focus series on value added assessment. General policy implications are discussed as it relates to the State of Ohio's push towards instituting a value added assessment model. The writer argues that a VAAS will allow teachers to focus more on individual student progress rather than a simple mean passing score for a group.

Relevancy to main topic: Any article written by a politician that is published in a journal such as JEBS is significant. It illustrates the political nature of the VAAS.

10. Chappuis, S., Stiggins, R.J., Arter, J. & Chappuis, J. (2004). *Assessment FOR learning: An action guide for school leaders*. Portland, OR: Assessment Training Institute.

This was one of the first books published by the group from the Assessment Training Institute. It is divided into four parts and can serve as a professional development tool to help educators and planners begin to increase the assessment literacy of teachers.

Relevancy to main topic: Merely having an assessment/accountability system doesn't ensure teachers will instantly know how to look at data and determine how they can alter their instructional practices. This can serve as a professional development tool for educators looking to increase the assessment literacy of teachers. In order to differentiate instruction properly, educators must first understand assessment and data analysis.

11. Chubb, J.E. & Moe, T.M. (1990). *Politics markets America's schools*. Washington, D.C.: The Brookings Institution.

This comprehensive study by Chubb and Moe is an often cited piece of research that analyzes the organizational structure of the public school system as it relates to a number of different output variables. In the book, the authors criticize the public school accountability system indicating that while some of the new strategies to determine student achievement might excite researchers in colleges of education, it seldom finds its way into the politics of decision making. “People look at a school’s average test scores and jump to conclusions—unwarranted conclusions—about school performance, with far-reaching consequences of their assessments of problems and solutions (p. 198)”.

While the authors admit that testing requirements make good sense, they believe that they have little chance of improving schools in any significant way. In the end, school choice is the suggested reform strategy by these authors.

Relevancy to main topic: Criticizes the accountability systems of schools because they rely on average test scores. The VAAS system would apply the same criticism and offers a new way to measure student achievement.

12. Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D. & York, R.L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

The report by Coleman and his colleagues that was submitted had a profound impact on the desegregation of schools. In sum, using an extremely large sample size, Coleman and his colleagues concluded that the achievement of children had very little to do with schools and the quality of instruction. Achievement was more strongly related to

home, neighborhood and peer interactions. This hypothesis virtually negated the impact of quality instruction.

Relevancy to main topic: A historic reference that serves as a baseline of how far the field has come with relationship to statistical regression and the creation of effect size. We now know that teachers and schools have a very large effect size on the achievement of children.

13. Crundwell, R.M. (2005). Alternative strategies for large scale student assessment in Canada: Is value-added assessment one possible answer. *Canadian Journal of Educational Administration and Policy*, 41, 1-21.

This article quickly reviews the state of educational accountability in the U.S. and draws four parallel issues existing in Canada. These include increased pressure to show what is working in schools; public demands for assessment tools that provide evidence that students are meeting standards; communication of test scores to the public and the psychometric properties of the assessments themselves as well as how the data is used to draw accountability conclusions. The author discusses the fact that at the time of the publication, there was no report of validity data for the assessments used in Ontario. Coincidentally these are the same conditions in Ohio. Next the authors address those non-instructional variables that may very well have a negative impact on the students' achievement performance.

Relevancy to main topic: Shows that the United States is not alone in its struggle to measure student achievement. Offers the value added method as a potential alternative to the current system of accountability.

14. Doran, H.C. (2003). Adding value to accountability. *Educational Leadership*, 61, (3), 55-59.

This source discusses the practical advantages of a value-added assessment system. The current system of educational accountability is criticized in that students are not randomly assigned to schools, failure to recognize growth towards proficiency is not reflected in most accountability models in that they rely on imprecise measures of cumulative achievement effects and many cut-score categories are not created using statistical techniques. A value-added assessment system, while complicated, will help the profession of education. It will encourage the use of student data in electronic format so that educators can analyze it much easier, there are sufficient sample sizes and using this method will allow test scores to be compared on a consistent test score metric through a process known as vertical equating. The criticisms of the value-added model are also discussed.

Relevancy to main topic: Related to the primary topic. How VAAS can help teachers make better instructional decisions.

15. Darling-Hammond, L. (2004). Standards, accountability, and school reform. *Teachers College Record* 106, (6), 1047-1085.

This article is a policy discussion and position statement. The author argues that the gathering of achievement data serves as an input to the accountability system however it is not the system itself. She analyzes initiatives in other states and districts that have successfully integrated an accountability system and points to investments in teacher quality and effectiveness as a key component. Specifically she discusses using standards and more authentic assessments to help leverage student achievement;

providing professional development opportunities that build teacher capacity to teach and assess student achievement in a more robust manner and encouraging collaborative relationships amongst educators.

Relevancy to main topic: The main points in this article are very important to keep in mind when developing or altering any accountability system. The data gathered doesn't serve as the accountability system and merely creating and implementing a more statistically reliable method to measure student growth (VAAS) does not ensure an increase in student achievement. While having teachers more critically look at data is important, an investment in their capacity to teach and formatively assess in a more critical manner is of equal importance.

16. DuFour, R. & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement*. Bloomington, IN: National Educational Service.

This is a contemporary book that has had a very large practical impact on school districts throughout the United States. It discusses a topic that has long been said to be critical in increasing student achievement. Much like Sergiovanni, DuFour and his colleague stress the importance of working together to improve. They argue that the focus of schools be on learning rather than teaching. School districts should have an articulated and *shared mission, vision* and a set of goals that are understood and embraced. This is then implemented by collaborative teams who work interdependently to achieve the district's and school's mission. Next, those that work within the professional learning community (PLC) engage in *collective inquiry*. Members of the PLC question the status quo and seek out new ways to improve. *Action orientation* and *experimentation* are encouraged in the PLC. Members try new things and evaluate the outcomes and make data-based decisions in a collaborative manner with a focus on

continuously improving previous practices. The final component of the PLC is its *results orientation*. Tangible results are the focus rather than suppositions.

Relevancy to main topic: While the concepts put forth in this book are not new to the field of education they are articulated well. DuFour doesn't offer a prescription of how it ought to be, but rather identifies some common elements that increase the likelihood of a district functioning more like a team. While creating a VAAS will provide educators in Ohio with something they have not had in the past, introducing it will not automatically create better results. Districts who embrace the concepts that DuFour discusses and begin to build the capacity to support them most likely have a higher probability of successfully implementing Ohio's proposed VAAS.

17. DuFour, R., DuFour, R., Eaker, R., & Karhanek, G. (2004). *Whatever it takes: How professional learning communities respond when kids don't learn*.

Bloomington, IN: National Educational Service.

This book builds on the previous work of DuFour and his colleagues. Within the Professional Learning Community (PLC) three fundamental questions are asked: What do we want students to know? How will we know when they've learned it? How does the PLC respond to students who experience difficulty learning that which we want them to know? The book uses practical examples and the experiences of other educators from five school districts to help educators begin to consider how a PLC functions.

Relevancy to main topic: Theory and ideas are stimulating to think about however if they can not be practically applied educators will not embrace them and implement. DuFour and his colleagues demonstrate how the ideas of a PLC can be implemented in different ways within five public school systems. The implications for VAAS are that in order to provide a higher probability of its success, a strategic effort to build the capacity of educators to work together in a format that embraces the concepts of a PLC is paramount.

18. Dusenbury, L, Brannigan, R., Hansen, W., Walsh, J. & Falco, M. (2005). Quality of implementation: Developing measures crucial to understanding diffusion of preventive interventions. *Health Education Research Theory and Practice*, 20 (3), 308-313.

This research article has the potential to become an often cited piece of literature with regards to program implementation. The authors reviewed those variables that previous research has shown to be of some importance when the discussion of program implementation by teachers is questioned. These variables included: 1) exposure to the program; 2) adhering to the methods and delivering the curriculum as it was intended; 3) active engagement by students receiving the program; 4) modification of the program by the teacher to meet the needs of the participants; 5) attitudes by the teachers implementing the program; 6) the degree of understanding that the teacher has with regards to the programs aims and objectives; 7) the teachers' prior experiences.

The purpose of this study was to conduct a quality check on the implementation of a drug abuse prevention program using an observational technique. In other words, teachers were trained on how to implement a program. They were then observed during a session to determine the degree to which they were implementing. The observers were two, Ph.D. researchers. The observers used a rating tool to measure adherence to the program, quality of the implementation and the engagement, how the teacher might have adapted the program and the teachers' attitudes as the program was being taught. The findings demonstrated that teachers varied in the degree of implementation. The researchers believe that program implementation can be measured through observation and interview as opposed to simple self-report by the teacher. All teachers made

adaptations to the program. Experienced teachers were found to adhere to the curriculum to a greater degree and deliver it in a way that was rated by the observers as more engaging to students. Experienced teachers also were found to better communicate the program goals and objectives.

Relevancy to main topic: While this article seems like somewhat of a stretch in that it does not relate to the topic of VAAS, it does nonetheless discuss some very important variables that all professional developers are interested in. The purpose of my dissertation will be to explore or illuminate the professional development repercussions of a VAAS. One of those very important questions will be how effective a trainer of trainers model might be given that Ohio's VAAS training model will be just that.

-
19. Ellett, C.D. & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, 17 (1), 101-128.

The purpose of this article was to provide a review of the research related to teacher effectiveness in the United States. The article focuses in on a review of the methods to chose and assess teachers along with the links between effective teaching and effective schools. The authors provide a historical review of the way in which teachers were evaluated and held accountable. They begin with the early part of the 1900s where teachers were chosen and considered effective based on personal characteristics. Teacher evaluation became more focused on a behaviorism model as teacher behaviors linked to student learning outcomes began to be studied. In the 1980s many university programs began to reform their teacher preparation programs. According to the authors, the evaluation and accountability movement began to take hold. In 1980, Georgia became the first state to evaluate teachers based on their on the job performance. Beginning in

the later part of 1980 and into the present, the teacher evaluation systems changed their focus from what was taught to what is learned. The focus on learning has continued to be the most important component.

The authors next provide a historical analysis of the research on school effectiveness and divide this analysis into 4 stages. During stage one, the conclusion was that schools have very little influence on student achievement. This was evidenced by report submitted by Coleman and his colleagues in 1966. Stage two began the study of school effectiveness research as it began to decipher the influence that schools have on student achievement. Student input was analyzed from both the teacher and classroom level and researchers began to observe student outputs based on the instructional practices that schools were engaging in. At the onset of stage three, the authors reported that school effectiveness research (SER) focused on an “equity orientation.” With this orientation, researchers tended to focus only on low socio-economic status (SES) districts rather than looking across the SES spectrum. This practice was criticized. The fourth stage expanded the focus of SER and it included and continues to include a wider array of study. More sophisticated statistical analyses have been developed in order to more accurately and fully study the school effect.

The article concludes with a summation of the strong relationship between effective teachers and effective schools. The authors cite and discuss Levine and Lezotte's research in 1990 that identified the characteristics of effective schools. This included: outstanding leadership, effective instruction, a focus on student learning target mastery, school climate and building culture, a well articulated student expectations,

frequent monitoring of student progress, parental involvement and job embedded staff development.

Relevancy to main topic: While a VAAS might be able to determine the output (student achievement) it essentially is a function of the impact that a school and a teacher have on a student. This article provides a very nice historical review of what the focus in education is; that being student achievement. A Value Added Assessment System is one of the ways in which to measure student achievement gain. The question will continue to remain what is a district to do with/to schools and teachers who are having difficulty adding value. The answer—investment in human capital.

20. Gutherie, J.W. (2005). An “education professions performance development act”: A prospectus for providing “highly qualified” and more motivated teachers and leaders for America’s schools. *Peabody Journal of Education*, 80 (3), 6-14.

This article is a policy review of the current state of the teacher and administrator appraisal process that exists in most states. The author criticizes the current teacher salary system and at the same time he explains why it is in existence. The author argues that people understand the current appraisal system however it far from useful. In an area of increased educational accountability the author believes that school performance is judged by the learning produced. Given this fact, the current teacher appraisal system is ineffective. He adds that in the past, it has been difficult to measure teacher and building effectiveness due to the many variables that are out of the control of teachers and administrators. This however, he argues, has begun to change as the result of the value-added approach. At the conclusion of the article, the author proposes an Educations Professions Performance Development Act. The proposed act has seven parts and it includes provisions that would authorize funding for a number of different things. These

include: determining methods to measure the value that teachers add to student achievement; experiments to determine if performance incentives for teachers and administrators make a difference in student achievement; measurement of teacher preparation programs; alternative teacher training programs; redefining No Child Left Behind's, Highly Qualified Teacher requirements; creating grant programs and redesigning administrator appraisal systems similar to those that would be devised for teachers.

Relevancy to main topic: Teacher and administrator effectiveness is at the core of a value added assessment system (VAAS). One could provide an argument that a method increase the probability of teachers and administrators adding value to student achievement might be through performance incentives. In addition, while VAAS is a measurement system, it is the system that Ohio's legislature has mandated and as a result there are educational policy and potential contractual issues associated with it. This article begins to suggest policy revisions to the ways in which administrators and teachers might be evaluated and suggests that a VAAS has the potential to help determine the impact that that teachers and administrators have on student achievement while controlling for variables beyond the their control.

21. Hammond, L.D. (2004). Standards, accountability, and school reform. *Teachers College Record*, 106 (6), 1047-1085.

This paper provides a comprehensive review of the school reform movement associated with standards based education and accountability. In the article, the author speaks to the different types of educational accountability and our push to try and create accountability through the use of testing. She argues that testing is the information obtained within the accountability system and it does not serve as the accountability system itself. The author suggests four strategies that could lead to improving learning.

These include an investment in professional development that focuses on preparing teachers to teach to a wider array of students; restructuring the design of the school to provide students with more individualized opportunities to learn; investing in formative assessments that are directed at providing teachers and students with feedback to how learning is progressing; and providing supports for students when they are in need. The author then provides a qualitative review of schools in Connecticut, New York City and New Haven, California. These three examples have a set of common characteristics in their reform efforts. The focus has been on an investment in human capital via professional development and supports for student learning.

Relevancy to main topic: The article criticizes the standards-movement from the standpoint that it is not enough to simply want high standards. In order to actually achieve this aim, there are some policy implications and best practices that need to occur. These include an investment in teachers through professional development.

22. Heck, R.H. (2000). Examining the impact of school quality on school outcomes and improvement: A value-added approach. *Educational Administration Quarterly*, 36, (4), 513-552.

This article is a quantitative study that included a subset of public school students from Hawaii. Its purpose was to evaluate a multilevel approach to using value added. The author first reviews three approaches that can be used to compare school outcomes. The first approach is the average performance of all the students. This method is criticized because it does not account for previous levels of learning and serves as a gross estimate of student achievement. Method two compares schools in similar comparison groups that are grouped based on similarities in pre-specified variables like SES. The problem with this approach is that sometimes, individual schools within comparison

districts vary significantly; particularly in larger districts. A third method looks to statistically “adjust the academic outcomes statistically for key indicators” that have shown to have an impact on student achievement. These adjustments are made prior to any comparisons in achievement performance.

This study used a value added model that made statistical adjustments to factors that have been shown to have an impact on student achievement that are not within the control of the schools. After doing so, the study explored the variables of school quality to see how they had an impact on student achievement over time. They found that schools who reported having high quality educational environments also report higher than expected levels of student achievement after the external variables were statistically controlled (adjusted). The author advocates a value added model that statistically adjusts for differences between students prior to determining that schools and teachers add.

Relevancy to main topic: The author reviews the effects that schools, teachers and communities can have on student achievement. However, when comparing this model to that which William Sanders created, it is much different. The TVAAS model does not include controls for student level demographics because to do so would admit there we would expect to observe different potential levels of achievement when certain variables are considered (race for example). The TVAAS model uses at least 3 prior student specific data points in order to consider student demographics. The model proposed by this author does not. There in lies the differences between these two value added models. A very important difference.

23. Hershberg, T. (2005). Value-added assessment and systemic reform: A response to the challenge of human capital development. *Phi Delta Kappan*, 84 (4), 276-283.

This article reviews the political compromise that needed to be reached in order to have NCLB pass, but calls for many changes to the current way in which the public school operates. Hershberg puts forth an argument by indicating that previously public school was set up to do three things: make people literate, socialize a population that was very diverse into an American set of values and identify the brightest students so that they can go to college and lead the country. These were the goals of education within an industrialized era. These goals and this view needs to be fundamentally changed if we are to compete in a global economy. It will require an investment in human capital with the goal of having all students meet high standards. He reviews the impact that NCLB has had on the public school and puts forth an argument that if we are to achieve our goal of educating all students to high standards we will need more leverage than that which NCLB has offered. We need to move from measuring achievement to measuring growth towards the achievement of standards.

Relevancy to main topic: This article does a nice job of framing the struggle that faces public education. How can we improve upon NCLB without abandoning the idea that we expect all students to reach proficiency or at least work towards it. The author proposes a value-added growth model and specifically mentions the Sanders system that Tennessee, Pennsylvania and Ohio have adopted as potential models.

24. Hershberg, T., Simon, V. & Lea-Kruger, B. (2004). The revelations of value-added. *School Administrator*, 61 (11), 10-12.

This article reviews how a value added assessment system can have a practical impact on student achievement. Three principals discuss how the value added assessment model has focused them and their teachers in helping students achieve. Arguments are provided why simply measuring achievement as opposed to student growth serves as an inefficient and insufficient manner to determine how schools are helping students. The authors point out how globalization has changed the kind of education needed for America to continue to remain competitive.

Relevancy to main topic: Points out the practical impact that a value added assessment system might have on teaching and learning through real examples and testimonial. Points out how NCLB will not be enough to stimulate real educational reform and that using an assessment system such as value added will help facilitate reform efforts.

25. Hill, P.W. (1998). Shaking the foundations: Research driven school reform. *School Effectiveness and School Improvement*, 9 (4), 419-436.

This article discusses a new model for educational effectiveness that employs a research driven school reform model. While the author writes about schools in Australia, there are many similarities in the reform efforts taking place in the U.S. The author argues that previous reform efforts “failed to penetrate beyond the classroom door to eliminate ineffective classroom practices (p. 420).” The author cites three reasons why the previous research design to study school effectiveness needs to change. First, the research study models have been top-down in nature and driven by bureaucracy. Researchers need to make connections to where the learning takes place—the classroom.

Second, the focus has been too narrow in that it has not focused on specific interventions or curricular reform initiatives. Finally, the research focuses too much on correlation and not enough on causation. The author cites previous research related to the impact that effective schools and effective classes can have on student achievement and calls for longitudinal studies to measure the impact that schools and teachers have on student achievement levels. The new paradigm will have a 3-tiered focus in that it will look at the System (state and entire district), the school (the individual places of learning) and the Innovation Network (the programs and interventions).

Relevancy to main topic: This article focuses on how reform might be measured and discusses professional development implications. Given that these are the same issue in another country, it shows that we can learn much by collaborating with those experiencing the same dilemma as.

26. Jamentz, K. (2002). *Isolation is the enemy of improvement: Instructional leadership to support standards-based practice*. San Francisco: WestEd.

This book serves as a resource for district professional developers. In order for students to reach high standards, teachers must begin to operate in a different manner. The author reviews the traditional model of teaching that incorporated a very frontal approach that didn't account for learner needs to one that is more student centered in that merely introducing a topic and talking about it results in immediate learning. Teachers must continue to check in with their students and re-teach as necessary.

Instructional vignettes are provided to share with teachers in a professional development setting in order to demonstrate how the concept of re-teacher and collaboration should work. Along with providing some very practical materials that staff

developers can use, the main position of the author is that teachers need to collaborate with one another much more so than has been typical in the past if we are to ensure high levels of learning for all.

Relevancy to main topic: The philosophy and theory presented in this professional development resource is invaluable. A VAAS system will identify which students need intervention, however it will not provide teachers with the professional development that will allow them to be effective. This resource will help professional developers prepare for ways in which they can improve teachers' instructional practices.

27. Jepsen, C. (2005). Teacher characteristics and student achievement: Evidence from teacher surveys. *Journal of Urban Economics*, 57, 302-319.

This study utilizes data from a warehouse called Prospects which is part of the Title I program. Utilizing this data, the researcher analyzed data sets of a first and third grade cohort. In addition, questionnaires were given to the students, parents, teachers, and principals. Data points were collected for the first and third grade cohorts in 1991, 1992, 1993 and 1994 using the Comprehensive Test of Basic Skills—4th edition. The researcher indicated that gathering data from lower stakes tests provided an advantage when “compared to state-specific high-stakes tests where teachers have incentive to teach to the test (p. 304).” Ultimately, the researcher concludes that measuring teacher attribute effects on student achievement is difficult. Classroom effects were analyzed and said to contribute to 25 to 40 percent of the variation within student achievement levels suggesting that classrooms make a difference. The author attributes these to both peer and teacher effects. While classrooms make a difference according to this author, he reported that there were few links between the observed teacher characteristics and student achievement. Those observed teacher characteristics included: class size, teacher

experience, degree obtained by the teacher, certification in the area the teacher teaches, presence of an aide in the classroom, use of computers, hours of homework teacher enthusiasm and adequacy of teaching materials.

My opinion of this study is that it was poorly reported. I believe that this has much to do with the use of a national database. There are many studies that the author could have used that provide sufficient evidence of the kinds of instructional characteristics employed by teacher make large differences and these are in fact observable. Simply looking at certification, use of a computer, adequate materials, etc., is a sign of ineptness. A quick review of the literature demonstrates that teachers who have clear learning targets and articulate those to students, frequently and formatively assess student learning levels and provide feedback to them are important teacher attributes that have an impact on achievement. This researcher not only didn't use them, he didn't talk about them and the study was from 2005.

Other parts of his study are questionable as well. The statement of low stakes tests having an advantage over high stakes ones because teachers teach to the test is troublesome. Teachers teach the learning targets prescribed by the content standards and the state mandated tests measure student obtainment of those standards. I would be critical of the way in which the data were analyzed as well. Four data points were collected on each student, yet a value-added measurement system wasn't incorporated to measure the obtained score from the predicted score.

Relevancy to main topic: A study that claims classrooms have an impact on student achievement is important. However, there is a sophisticated and a rather unsophisticated way to go about measuring and I believe this researcher engaged in the former.

28. Keeves, J.P., Njora, H. & Afrassa, T. (2005). Measuring value added effects across schools: Should schools be compared in performance? *Studies in Educational Evaluation, 31*, 247-266.

This was another study that investigated the different value added models that exist. Subjects were from different schools, but from the same school system and consisted of two cohorts. An HLM was used with different multilevel models much like the Tekwe et al article in this bibliography. The results reported were that depending upon the model and variables used, very different results were found with regards to differences between the schools.

Relevancy to main topic: This study provides further evidence of the difficulties associated with the different value-added models that exist. A model that accounts for student differences prior to the analysis and one that gathers at least 3 data points prior to the analysis as a method to “block out” student differences will get different results. This is controversial and more research is needed in order to determine which is the best model to use.

29. Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value added assessment system. *Educational Evaluation and Policy Analysis, 25* (3), 287-298.

After providing an overview of how the Tennessee's Value Added Assessment System (TVAAS) system works, this article presents a critique of it. One of the critiques centers on the definition of teacher effectiveness—specifically the notion that students who are assigned to teachers exhibiting greater achievement gains are viewed as being with a teacher who is more effective. This researcher is critical of Sanders in that he believes that “real evidence” should be offered to measure teacher effectiveness. This real evidence would include using student SES or other student characteristics as

predictors of achievement gains. Kupermintz asserts that the TVAAS model incorrectly assumes that teacher effects are “independent, additive and linear (p. 290).” He believes that teaching and learning is an interactive phenomenon and that a model such as TVAAS does not capture this

The TVAAS model is a value added model that does not seek to control for covariates such as student background. The idea is that because multiple data points are for individual students, students serve as their own baseline and that by using at least 3 data points any specific, individual student differences were accounted for already across the different measures—this is called “blocking out.” Kupermintz is critical of the “blocking out” method and states that research by Ladd and Walsh (2002) found large correlations between SES and a school's value added scores. In other words students from low SES experience lower academic gains than those from high SES. Other value added models look to account for these covariates. One fine problem with the TVAAS model that is addressed is the assignment of students to teachers. Typically, students are not randomly assigned to teachers and at times the classroom makeup of students can differ dramatically. These peer effects can have an influence on student gains. Kupermintz concludes with a call for Sanders to release the data that he has to the research community so that it can be independently studied.

Relevancy to main topic: The article is a direct critique of the TVAAS model created by Sanders. While it indicates that value added models have promise, it calls for an extensive validity study before using it as an accountability system.

30. Lewis, M.S. & Ruhil, A.V.S. (2006). Achievement gains study of SOAR pilot and match districts. Athens: Ohio University, The Voinovich Center for Leadership and Public Affairs.

This report of was written by two senior research associates from Ohio University. By using a quasi-experimental design, the research study sought out to answer two questions: 1) Was the change in Ohio Proficiency Test performance from 2002-03 to 2004-05 greater for Schools' Online Achievement Reports (SOAR) districts as compared to non-SOAR districts; and 2) Which, if any, academic areas (reading, math, citizenship, science) showed the greatest change? SOAR districts are divided into four cohorts using an ordinal scale. Those with the lower numbers have been implementing the value-added pilot the longest amount of time (e.g., SOAR I districts were involved at the onset followed by SOAR II and so on). This study used the 63 SOAR districts from cohorts I and II (i.e., SOAR I and SOAR II). While completing this project, the researchers reported they found varying levels of SOAR implementation as revealed by a questionnaire completed by the district superintendent. While some districts did take part in the study, the district staffs that were obtaining the professional development either did not train staff properly, were prevented from implementing the project or perhaps a combination of the two occurred. Nonetheless, their rigor of implementation was not as great as compared to full SOAR implementers. Full SOAR implementers were described by the superintendent as using data at the building level to inform decisions about student instruction. There were fourteen full SOAR implementers identified.

This project used four designs in order to address the research questions. In Design 1, SOAR districts were matched with 4 unique match districts using ODE's

Similar District Comparison matching. In Design 2, the 63 SOAR districts were matched to 52 of the best propensity score match districts. The authors described what propensity score estimation entails. It is too long to describe here. In Design 3, the 14 SOAR full implementers were compared to their propensity score match districts. Design 4 entailed a longitudinal subset comparison between students with 2003 4th grade proficiency scores and 2005 6th grade scores.

In their conclusions, the researchers indicated differences between SOAR and non-SOAR districts were not evident; however differences did become evident when the SOAR full implementers were investigated. A major point made by the authors was the fact that having access to value-added data does not guarantee that a district will make use of it to shape district, building, teacher and student growth plans. The authors also indicated that their absence of randomized design and controls clouds their quasi-experimental design. However, any information that can be gleaned from the 14 SOAR full implementers would be valuable for Ohio's statewide rollout of value added. This study is directly related to that which I intend to do and the researchers indicated that there appears to be a need to follow up with the SOAR full implementing districts.

Relevancy to main topic:

-
31. Mahoney, J.W. (2004). Why add value in assessment? *School Administrator*, 61 (11), 16-18.

This is a very practical article related to the components of Ohio's value added assessment system written by the individual who is the executive director of Battelle for

Kids. While Dr. Mahoney may have some bias as towards which value added might be better than the other, the article provides a good overview of the purpose of such a model.

Relevancy to main topic: This article was written by the person who is responsible for overseeing the implementation of Ohio's EVAAS. It quickly reviews the data that has been gathered thus far by Ohio's school districts who are participating in the SOAR project.

32. Marzano, R.J. (2003). What works in schools: Translating research into action.

Alexandria, VA: ASCD.

This book presented by Marzano reviews the meta-analysis that he conducted. He puts forth 3 major assertions. First, the studies that previously indicated schools do not significantly affect student achievement have been misinterpreted and when they are interpreted correctly they support the impact of schools on students. Second, the research on effective schools shows they have a positive impact on student achievement. Schools that are highly effective produce results that almost entirely overcome the effects of student background. Marzano reviews school level factors, teacher level factors and student level factors that he purports have a major impact on student achievement.

Relevancy to main topic: Uses a meta analysis technique to identify teaching practices that research has shown to lead to greater achievement.

33. McCaffrey, D.F., Lockwood, J.R., Koretz, D.M. & Hamilton, L.S. (2003).

Evaluating value-added models of teacher accountability. Santa Monica, CA:

RAND Corporation for the Carnegie Corporation of New York.

This technical report is comprehensive in that it reviews the general concepts associated with value-added models, provides a literature review of the three main value-

added models and discusses modeling longitudinal data to estimate teacher effectiveness. Seven recommendations for future research are given and include: Developing databases that can support value added modeling (VAM) estimation of teacher effects; Develop computational tools for fitting VAM; Link VAM teacher-effect estimates to alternative measures of teacher effectiveness; Empirically evaluate the potential sources of errors that were identified in the report; Estimate the prevalence of factors that contribute to the sensitivity of teacher-effect estimates; Incorporate decision theory into VAM and use research and auxiliary data to inform modeling choices.

Relevancy to main topic: Discusses the different value added assessment models that exists and provides insight as to the advantages and disadvantages of each.

34. McCaffrey, D.F., Lockwood, J.R., Koretz, D., Louis, T.A., & Hamilton, L. (2004).

Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29 (1), 67-101.

This article was much like the 2003 article that was written by this set of researchers. The difference is it appeared in a special edition of JEBS devoted to value-added assessment. The purpose of the model was to specifically look at the two competing value added models. One that uses a blocking technique to control for covariates (TVAAS) and the second that looks to specifically and statistically account for individual student differences before conducting an analysis of student achievement gains. The authors run a complicated simulation to test both models. They conclude that a value added model that uses the blocking technique in order to account for student variation (such as TVAAS) are “robust” if the sample of students is similar in makeup (i.e. they are a similar population of students). However, this model is not as accurate

when making comparisons between two schools or districts which have a significantly different population of students

Relevancy to main topic: This is one of the leading groups of researchers who have spent considerable time trying to evaluate the effectiveness of the different value added models. The main point that he made, comparing different populations of students and the reliability of the model is not a view that the creators of TVAAS would agree with. However, value added models and the implementation of them are political in nature. Ohio may very well look to use their comparison groupings when comparing the effectiveness of one district compared to another when it comes to adding value.

35. Morgan, J.G. (2002). Multiple choices: Testing students in Tennessee. Retrieved June 24, 2006, from <http://www.comptroller.state.tn.us/orea/reports>

This is an educational accountability report that was submitted to both branches of Tennessee's state government. The report contains information and recommendations that policy makers might use to improve Tennessee's testing program. The report contained some useful points. First, the report indicated that too few schools use test data to improve student learning. There was reference to confusing data reports and a lack of professional development provided to the educators who would be the ones responsible for reading and interpreting the reports. There were five administrative recommendations cited in the report. The second and perhaps most important to my dissertation is Tennessee's Department of Education needs to provide more professional development to schools and systems to help them learn how to use testing data in a meaningful way so that it can have a positive effect on student achievement.

Relevancy to main topic: This is Tennessee's self evaluation of the implementation of their model. Given that Ohio is attempting to implement a very similar VAAS, the information contained in this report would offer some insight.

36. Raudenbush, S.W. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, 29 (1), 121-129.

This article discussed and reviewed the fact that there are two causal effects in educational accountability systems; Type A and Type B. Type A Effects are of interest to parents and Type B are good for schools. A Type A Effect is the difference between a child's potential outcome in school 1 and that same child's potential outcome in school 2.

A Type B Effect is the difference between a child's potential outcome in school 1 when school practice P_j is in operation and that same child's potential outcomes in school 1 engages in school practice P_i .

The implications for a value added model (VAM) are explained. The Type B analysis would try to separate the practice at the school level and the teacher level, however both practices are not observable and are inaccessible in an accountability analysis. Mean differences in schools are simply differences in aggregate teacher effects. As a result, the author argues that VAM are best aimed at assessing the combined effects of context and practice at the classroom and building level.

Relevancy to main topic: A very technical article that serves the purpose of exploring the statistical complexities of a VAAS by one of the main researchers in the field of statistical analysis.

37. Reeves, D. (2002). *Making standards work: How to implement standards-based assessments in the classroom, school and district*. Denver, CO: Advanced Learning Press.

This was one of the first books written by Dr. Reeves. In the latter part of the 1990 there was a major push for standards-based reform. This book serves as a resource to help districts understand how and what it can do to help bring about this type of reform in the school district.

Relevancy to main topic: Relates to the professional development that might be provided to teachers so that they can become more adept at standards-based education.

38. Reeves, D. (2002). *The leader's guide to standards: A blueprint for educational equity and excellence*. San Francisco: Jossey-Bass.

This book written by Dr. Reeves serves as another resource to school leaders trying to bring about standards-based reform within their respective school districts. The book is divided into three sections. The first section explains the importance of standards and how standards are different than norms. Part two of the book addresses the kind of leadership that is needed in order to bring about change. The final section of the book has chapters devoted to the educational leaders within a state—specifically the superintendent, board members and state superintendent.

Relevancy to main topic: Relates to the professional development that might be provided to teachers so that they can become more adept at standards-based education.

39. Reckase, M.D. (2004). The real world is more complicated than we would like. *Journal of Educational and Behavioral Statistics*, 29 (1), 117-120.

This article questions the use of a simple conception of student growth in academic achievement. The author puts forth an argument that achievement is non-linear and making simple comparisons from grade to grade is not as simple as it looks. For instance, measuring a student's performance in one grade and making a prediction on what his or her performance might be on the next grade is dependent upon the kind of information assessed. A key point is made by stating "a mismatch between instruction and the assessment will result in an underestimate in students' change in performance." This would occur regardless of controlling for variables such as race and SES. The author cautions all who would use a value added model by indicating that we require students to know different things at different grade levels and the knowledge moves from facts and understanding to application and problem-solving. He calls for examining the assessments in great detail.

Relevancy to main topic: When I read this article I thought of the two value added "camps" that exist. This researcher appears objective in that he doesn't discuss whether or not it is proper or improper to control for variables. Rather, this researcher is looking at the knowledge that the tests themselves measure and cautions value-added researchers when making simple comparisons of growth from year to year.

40. Reynolds, D., Muijs, D. & Treharne, D. (2003). Teacher evaluation and teacher effectiveness in the United Kingdom. *Journal of Personnel Evaluation in Education* 17 (1), 83-100.

The authors in this article provide a historical overview of the research related to teacher effectiveness and teacher evaluation in the United Kingdom. This article

provides much of the same overview that one that looked at these issues in the United States. There appears to have been much research related to school effectiveness in the U.K over the past 30 years and very little devoted to actual teacher effectiveness of what kinds of instructional techniques or teacher variables lead to increased gains in student achievement. The article calls for more research and development related to reliable value added measures that are capable of measuring teacher techniques that have a causation effect on student achievement gains.

Relevancy to main topic: This article demonstrates that the struggle to determine teacher effects exists abroad as well.

41. Rubin, D.B., Stuart, E.A. & Zanutto, E.L. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and Behavioral Statistics*, 29 (1), 103-116.

This article takes a different view when analyzing value added models. The fundamental question the authors address is whether or not value added models are causal or descriptive. The authors in this study have an obvious expertise in causal-comparative statistics and point out the flaws of all of the proposed value added assessment models. While each model claims to be able to indicate that teacher A caused student B's learning gains as demonstrated by a gain from baseline, these researchers point out that such a comparison can't be made due to the lack of randomization of schools to districts, teachers to schools and students to teachers. In addition, these authors point out the problems of missing data in value-added models as well.

These authors offer another model that is more causal. They would argue that without randomization, attributing gains in student achievement to a teacher's effectiveness would be erroneous or would, at least, not tell the entire story. Their model would include three fundamental concepts of causal research—units, treatment and outcomes. In this case the units would be the schools or teacher. The treatment would be the kind of reward structure that would be incorporated into the district as a result of observed achievement gains or lack thereof and the outcomes would be the gain achievement scores themselves. In this sense a researcher here would be determining the causal relationship of implementing a particular value-added model. Does the reward structure cause gains in student achievement? The researchers indicate that in order to test out such a model you would need randomization, however this model would be more causal than a value added system. In the end the authors indicate that value added models should be considered as providing descriptive measures rather than causal effects.

Relevancy to main topic: This article touched on several valid points and made me consider things that I had not prior to reading it (namely causal vs. descriptive of VAAS). While I understand the point of a VAAS not be causal, their argument is nonetheless more theoretical than practical. Students will never be randomly assigned to teachers or schools.

42. Sanders, W.L. & Horn, S.P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education* 12 (3), 247-256.

This article was written by the founder of the Tennessee Value Added Project. The publication A Nation at Risk and then President George Bush Sr.'s Education Summit

with the nations governors, resulted in Tennessee Governor Lamar Alexander to create the Comprehensive Education Reform Act in 1989. In 1992, Tennessee's Educational Improvement Act created a strong teacher accountability system. The system determines the effectiveness of school districts, individual schools and teacher performance within the school based on the analysis of student achievement data over time. Research conducted from the massive TVAAS database has demonstrated that race, socio-economic status, class size and class heterogeneity are poor predictors of student academic growth. The effectiveness of the teacher was the major factor in determining student academic gains. The authors conclude by stating the need to gather annual "fresh" data regardless of whether or not it is standardized or criterion-referenced in order to more accurately determine the individual growth of students.

Relevancy to main topic: This article was co-written by the founder of the VAAS and it provides a historical reference as to how and why the VAAS was created in Tennessee.

43. Schacter, J. & Thum, Y.M. (2004). Paying for high- and low-quality teaching.

Economics of Education Review, 23, 411-430.

This study sought to begin to more clearly define what high quality teaching consists of. The researchers used three bodies of research on teaching to create a set of standards and rubrics which were then reviewed by teachers, administrators and curriculum specialists. Fifty-two teachers participated in this study as volunteers. Along with the teachers, the students who were assigned to the teachers participated as well. There were a total of 910 students. Over a period of 9-months, each teacher was observed 8 times. Teachers were unaware that they were being evaluated as they were being observed. Observations were counterbalanced and visited the teachers at different times.

In order to ensure reliability in the report from observers, observers were trained using videos of teachers. Teachers were rated on the 12 teaching performance standards that were created by teachers, administrators and curriculum specialists. These included: teacher content knowledge, lesson objectives, presentation, lesson structure and pacing, activities, feedback, questions, thinking, grouping students, motivating students, classroom environment, and teacher knowledge of students. Those observed were given a performance level of exemplary, proficient and ineffective. Narratives were included as well.

Results indicate that teachers who implemented effective teaching as evidenced by adhering to the 12 teaching performance standards had students who made considerable achievement gains compared to those that did not (effect size at .91). The researchers also found evidence that classroom composition effects quality teaching. In sum, the researchers reported that teaching quality as measured by their rubric and classroom composition accounted for 84% of the variance in student achievement gains in reading, math and language.

Relevance to main topic: Confirms previous studies which have indicated that teacher quality makes an enormous impact on student achievement levels. While creating an EVAAS will measure student learning gains, it can not identify quality teachers. Assuming that a quality teacher would get the most gain from students, one would need to observe such a teacher and identify whether or not the teacher engages in quality instructional practices. This study provides characteristics that teachers engage in who are said to exhibit quality teaching.

44. Schlechty, P.C. (1997). *Inventing better schools: An action plan for educational reform*. San Francisco: Jossey-Bass.

Much like the work of DuFour and his colleagues, Schlechty is a practitioner and his ideas related to educational reform reflect this. Schlechty does not provide a blue print that outlines a lock step reform design. Rather, his focus is on system variables and an effort to make the conditions right so that an educational reform effort can be sustained. Schlechty argues that student engagement in learning is important and that engagement and entertainment are not synonymous. Students need to be invested in what they do. In addition, like many reform effort models reviewed in the 1990's, a focus on clearly articulated district vision, mission and values is critical. This should be developed and embraced by the stakeholders throughout the system. As a result, rather than bringing in outside gurus to help the system reform itself, districts should build the capacity within. At the core of Schlechty's reform is the idea that teachers should stop working on students and start working on the work that they provide to them. The work that students engage in should be focused and clearly related to standards. Student work should encourage collaboration and cooperation. It should be novel. With a collective belief in the mission and vision along with student work that is engaging and focused another key component in Schlechty's reform effort is measuring results.

Relevancy to main topic: Schelechty is another contemporary author who, like DuFour, discusses a reform effort that is not meant to serve as a lock-step plan, but rather a collection of guiding principals. The main topic studied in this dissertation will be the VAAS. Introducing such a model in the absence of many of the concepts discussed by Schelechty may very well result in another failed reform effort.

45. Stiggins, R.J. & Chappuis, S. (2005). Putting testing in perspective: It's for learning. *Principal Leadership (Middle School Edition)*, 6 (2), 17-20.

The authors review the differences between two different kinds of assessment—formative and summative. Summative assessment occurs at the end of learning and is used primarily for accountability purposes. It is less frequent than formative assessment. Formative assessment is conducted during the learning process and is used by the teacher and student in order to monitor student progress towards learning goals.

Stiggins and Chappuis identify three approaches to making formative assessment more useful to teachers. The first approach is to test more often using summative assessments so that they can be used in a formative manner. The second approach is to design a data management system to allow educators to look at assessment data that has been gathered in a more efficient way so that they can improve their teaching practices. The third approach is engage in assessment for learning which entails a more frequent gathering of student evidence of growth towards specified learning targets. The authors point out the third approach is different than the first two for it not only provides teachers with feedback as to how their students are doing so that they can adjust instruction during the learning process, but it does the same for students. This is a key component to Stiggins' work and how it differentiates itself from others in the field.

Relevancy to main topic: EVAAS will create a system that will allow teachers to look at student performance on summative tests and it will be used as part of an accountability system to measure teacher effectiveness. Given this fact, teachers will need to learn how to formatively assess more frequently to not only improve their teaching practices, but to help students realize where they are with relationship to the achievement of the learning targets.

46. Stiggins, R.J., Arter, J.A., Chappuis, J. & Chappuis, S. (2004). *Classroom assessment for student learning: Doing it right—using it well*. Portland, OR: Assessment Training Institute, Inc.

This is an invaluable professional development resource that can be used as part of a district's effort to build staff capacity to engage in effective assessment practices. Stiggins and his colleagues build a convincing argument as to what constitutes a quality assessment system; specifically, 4 key questions are asked when constructing an assessment. First, why are we assessing (purpose)? Second, what are we assessing (which learning targets)? Third, how are we assessing (method...does it match the complexity of the learning target)? Fourth, how are we communicating to students their progress? This book provides many hands on professional development activities that professional developers can use with teachers as a means to increase staff capacity to more effectively use data to guide their instruction.

Relevancy to main topic: Relates to the professional development that might be provided to teachers so that they can become more adept at standards-based education.

47. Stone, J.E. (1999). Value added assessment: An accountability revolution. In M. Kanstoroom and C.E. Finn, Jr. (Eds.), *Better teachers, better schools*. Washington, DC: Thomas B. Fordham Foundation.

This chapter located in a book about how teachers and schools can become better outlines the advantages of Tennessee's Value Added Assessment System (TVAAS). 1) Effectiveness of teachers, schools and districts is measured by comparing students' current performances with previous ones. This allows the accountability system to look at gains in student achievement as opposed to simply looking at a student's position with

relationship to an arbitrary cutoff; 2) the statistically nature of the VAAS controls for individual student differences; 3) it can isolate the achievement effects and contribute them to individual teachers more accurately; 4) given its statistical foundation, it can account for student achievement despite missing or incomplete data and 6) it has the potential to compare gains in student achievement on a national level.

48. Stufflebeam, D.L. (2001). Evaluation models. Western Michigan University Evaluation Center.

This source is a ninety-eight page monograph that describes twenty-two approaches that are often used to evaluate programs. Included in this evaluation is the value-added model. The author believes that a central advantage of this type of evaluation is the systematization and institutionalization of a database of outcomes that can be used over time to study trends and improve performance. Students are used as their own controls and rather than focusing on static cut scores, the model emphasizes continuous improvement for all students. He indicates that such an evaluation system requires a well financed state education department and system-wide buy in. In addition, a computerized baseline of relevant input and output data is needed along with a powerful computer system to run the large-scale statistical analyses. Disadvantages of the model include the political volatility of it and its reliance on quantitative methodology and lack of balance with qualitative measures. The author acknowledges that while advances have been made with objective measures and hierarchical mixed models of assessment to determine the effects a system or person might have, no system can account for so many

complex variables and interactions that are responsible for the academic progress of students.

49. Tekwe, C.D., Carter, R.L., Ma, C.X., Algina, J., Lucas, M.E., Roth, J., Ariet, M., Fisher, T., & Resnick, M. An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29 (1), 11-36.

In this study the researchers sought out to evaluate three different value added assessment models of student achievement. These included a layered mixed effects model (LMEM), a hierarchical linear model (UHLMM) and a simple fixed effect model (SFEM). Using data from schools in a single Florida school district, the researchers evaluated the different models. Based on the results, the researchers recommended that the SFEM model should be preferred over the LMEM due to its complexity. They base this assumption on how highly correlated the models were with their findings in their analysis of student data ($r > .91$). The LMEM model is that which William Sanders and his colleagues have developed. The next finding was that the authors concluded they could not recommend the SFEM model over the AHLMM model. The differences between these two models is significant and it is one of the main variables that has put advocates of a value-added approach to student achievement into two separate camps. One model does not statistically adjust for what are referred to as “significant sociodemographic factors (p. 31)” whereas the other model (AHLMM) does. The researchers point out the main problem that those in the SFEM/LMEM camp have with those in the AHLMM camp. That is, making statistical adjustments for

sociodemographic variables has the potential to institutionalize lower expectations for the economically poor and racial minority population.

Relevancy to main topic: This article was a good one. It specifically addresses the issues in the two competing value added models and points out the potential issues with each.

50. Wong, K.K. & Nicotera, A.C. (2004). Brown v. board of education and the Coleman report: Social science research and the debate on educational equality. *Peabody Journal of Education*, 79 (2), 122-135.

This article reviews the impacts of the Brown v. Board of Education (1954) decision and the Coleman Report with specific emphasis on their long-term impact of social science research and how it has shaped the school opportunities for children. A quote from the Coleman Report stated

Differences in school facilities and curriculum, which are the major variables by which attempts are made to improve schools, are so little related to differences in achievement levels of students that, with few exceptions, their effects fail to appear even in a survey of this magnitude. (Coleman, 1966, p. 316).

While the authors acknowledge the validity of the findings of the Coleman Report, their point is impact that it had on educational policy decisions; particularly with reference to the above quote. This article reviews the impact and encourages social science researchers to be innovative and rigorous with relationship to their research methods in order to “push the notion of equality to new levels.”

51. Wright, S.P., Horn, S.P. & Sanders, W.L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57-67.

This was one of the first studies conducted in Tennessee to evaluate the effects of the state's VAAS. Subsets of data from Tennessee's annual state achievement test were analyzed from 1994 and 1995. One set of data contained 30 East Tennessee school systems whereas the other was from 24 Middle Tennessee systems. Student achievement gains were analyzed while considering the impact of class makeup, class size and student incoming achievement levels. Based on the results, the largest differences in student achievement levels were reported as being attributed to the classroom teacher's effectiveness as opposed to class makeup or class size. Higher achieving students were found to make somewhat lower relative academic gains when compared to average and lower achieving students.

-
52. Young, T.A. (1996). A study of the perceptions of Tennessee teachers, principals, superintendents, legislators, and department of education representatives regarding the impact of the Tennessee value-added assessment system.

Unpublished doctoral dissertation, The University of Alabama.

This dissertation study investigated the perceptions of those who made the determination to create a value added assessment system through legislation and policy and the perceptions of those educators affected by it. There were eleven research questions which were all aimed at measuring the perceptions of the participants through a survey research methodology. The instrument was designed by the doctoral student and

it consisted of thirteen items. The survey was field tested and had acceptable levels of reliability (.95). Two conclusions drawn by the author include:

- Perceptions show that most Tennessee educators (teachers, principals and superintendents) do not support TVAAS in its present application.
- Perceptions show that most Tennessee government officials support TVAAS.

Perhaps the most important points of the dissertation are relayed in a section called *Implications for Educational Decision Makers*. In this section, the author points out the need for all parties to collaborate with one another when creating a state accountability system such as VAAS. In addition, the author recommends that such a system should be piloted, should contain a structured orientation system and staff development program prior to its implementation and stresses the need for all educators to work together to identify educational effectiveness models.

Needed References

- Burns, M.K., Courtad, C.A., Hoffman, H, & Folger, W (2004). A comparison of district-level variables and state accountability test results for public elementary and middle schools. *Psychology and Education, 41* (2), 17-26.
- Madaus, G.F., Kellaghan, T., Rakow, E.A., & King, D. (1979). The sensitivity of measures of school effectiveness. *Harvard Educational Review, 49* (2), 207-230.
- Meyer, R. H. (1996). Value-added indicators of school performance, In E. A. Hanushek, and D. W. Jorgenson, (Eds.), *Improving America's schools: The role of incentives*, (pp.197-223). Washing, DC: National Academy Press.
- Raudenbush, S.W., & Willms, J.D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics, 20* (4), 307-335.
- Rosental, R., & Rubin, D.B. (1982). A simple general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74* (2), 166-169.
- Sanders, W. L. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education, 14*(4), 329-339.

References Not Used in Annotation, but Could Be in Dissertation

53. Webster, W.J. (1995). The connection between personnel evaluation and school evaluation. *Studies in Educational Evaluation, 21*, 227-254.

54. Webster, W.J. & Mendro, R.L. (1995). Evaluation for improved school level decision-making and productivity. *Studies in Educational Evaluation, 21*, 361-399.

55. Rodgers, T. (2005). Measuring value added in higher education. Do any of the recent experiences in secondary education in the United Kingdom suggest a way forward? *Quality Assurance in Education, 13* (2), 95-106.

56. Kupermitz, H. (2002). Value-Added assessment of teachers. In A. Molnar (Ed.), *School reform proposals: The research evidence (EPSL-0201-101-EPRU)*. Educational Policy Research Unit: Arizona State University.

Relevancy to main topic:

57. Schmoker, M. (1999). *Results: The key to continuous school improvement*. Alexandria, VA: Association for Supervision and Curriculum Development.

Relevancy to main topic:

58. Langer, G.M., Colton, A.B. & Goff, L.S. (2003). Collaborative analysis of student work: Improving teaching and learning. Alexandria, VA: ASCD.

Relevancy to main topic:

59. McCaffrey, D.F., Lockwood, J.R., Koretz, D., Louis, T. & Hamilton, L. (2004). Let's see more empirical studies on value-added modeling of teacher effects: A reply to Raudenbush, Rubin, Stuart and Zanutto, and Reckase. *Journal of Educational and Behavioral Statistics*, 29 (1), 139-143.

Relevancy to main topic:

60. Vaughan, A.C. (2002). Standards, accountability, and the determination of school success. *The Educational Forum* 66 (3), 206-213.
61. Schoneman, T. (2003). *Collaborating for high standards: Analyzing student work*. San Francisco: WestEd.

Relevancy to main topic: *Relates to the professional development that might be provided to teachers so that they can become more adept at standards-based education.*