

Post-ESSA VAM Scholar Survey Online Supplement C

Post-ESSA Survey Instrument for VAM Scholars

Email

Dear [Firstname] [Lastname],

You are being invited to participate in a research study to determine the opinions of scholars and researchers about the use of value-added models (VAMs) in teacher evaluation systems. You were selected for inclusion in this study based on your (co)authorship of the following article(s), [titles of identified articles], which has/have been published in (a) highly-esteemed, peer-reviewed journal(s).

This anonymous online survey should take about 20 minutes to complete. You may complete the survey on your computer, tablet, or mobile device, though we recommend taking it on a computer with a keyboard as the survey does contain some open-ended questions.

Please follow the appropriate link below.

Thank you for assisting us with this research,
[removed for peer review]

Follow this link to the Survey:
[survey link]

Follow the link to opt out of future emails:
[opt-out link]

Thank you for your interest in participating in this study. A team of researchers from [information removed for peer review] is interested in learning more about your opinions on the use of value-added models (VAMs) in teacher evaluation systems. This anonymous 20-minute online survey will ask you about your viewpoints on several aspects of VAMs and VAM use. This survey is mobile/tablet compatible; however, you might find it easier to take it on a computer with a keyboard as there are some open-ended items.

Survey

Introduction: You have been identified as a potential participant in this study based on your authorship of one or more published articles on the use of value-added models (VAMs) specific to teacher evaluation (see the email inviting you to participate in this study for the list). We are conducting an anonymous online survey to assess the opinions of VAM “experts” (i.e., scholars/researchers who have published articles on VAMs or VAM use in respected peer-reviewed journals between 2002-2016), especially in light of recent legislation (e.g., the Every Student Succeeds Act; ESSA, 2015).

Purpose: The purpose of this study is to better understand the opinions of identified VAM experts about VAMs and VAM use. As VAMs and education policy related to VAMs continue to be contentious and debated, both pragmatically and empirically, this study will continue to shed light on how experts in the field feel about VAMs and related VAM use.

Procedure: If you consent to participate in this study, you will be asked to complete an anonymous online survey consisting of approximately 50 close-ended items and 2 open-ended items related to VAMs and VAM use. The survey also provides 9 additional text boxes where you may provide additional comments if you wish to do so. All items are optional and you may skip any question.

Voluntary Participation: Participation in this study is completely voluntary. You are free to withdraw at any time. You may decide to skip any question or discontinue participation at any time without penalty. Deciding to participate or not will have no effect on your position at your institution/organization, your relationship with the researchers, or your relationship with [information removed for peer review] or [information removed for peer review].

Anonymity Protection: Although you were identified for inclusion in this study based on your scholarly publications, this survey is anonymous. The Qualtrics online survey system used for this study stores participants' contact information and survey responses in separate data files and offers an option to Anonymize Responses by removing personally identifiable data from the response record. Researchers intend to release the list of references from which potential participants were identified and to report on the response rate for the survey but will not release personally identifiable information about the authors who responded to the survey, including but not limited to name or institution/organization affiliation. Data from this survey will be stored on fully secured and password protected [information removed for peer review] and [information removed for peer review] servers. Only the research team will have access to the data and will only use the data for the purposes of the present study. Since this is an online electronic survey and some employers may use tracking software for computer use, you may want to complete this survey on a personal computer. If you use a public computer or a computer to which others may have access, you should not leave the survey open and unattended. You should also clear your browser cache after completing this survey. In addition, please do not include any personally identifiable information in your responses to open-ended items.

Risks: This study poses no more risk than is experienced during everyday life.

Contact Information: If you have any questions about this study or your participation in this survey, you can contact [information removed for peer review]. If you have questions about your rights as a research participant, you can contact the [information removed for peer review] Institutional Review Board (IRB) Chair at [information removed for peer review].

Do you consent to participate in this study?

Yes

No

Knowledge, Experience, and Perspective on Value-Added Models

For the article(s) referenced in your invitation email which of the following roles did you play on the research team? (*Please check all that apply. If multiple articles were referenced in your invitation email, check an item if you filled that role for any of the articles listed.*)

- Conceptualized (or helped conceptualize) the paper/study
- Wrote (all or part of) the paper
- Conducted literature review
- Authored survey, interview protocol, or other data collection instrument
- Conducted interviews, focus groups, or other qualitative data collection
- Created and analyzed simulated data
- Analyzed non-simulated data
- Analyzed qualitative data
- Other (please specify):

On a 5-point scale, how would you rate your professional expertise with value-added models (VAMs) including but not limited to your work on the publications listed in the email inviting you to participate in this study, and your research and scholarship, your experience, your education and training, and the entire body of your professional work?

- (5) **Leading Expert** - I am among the most knowledgeable on this topic, known in the field for my scholarly contributions.
- (4) **Expert** - VAMs are one of my primary areas of expertise; I have a great deal of knowledge on the subject, but there are other scholars that I consider to be the leading experts in this area.
- (3) **Experienced Scholar** - I have a great deal of knowledge and expertise about VAMs and have done substantial work in the area, but VAMs are not one of my primary areas of expertise (*or* my work in this area is not yet fully developed).
- (2) **Scholar** - I have enough knowledge on the topic of VAMs to feel comfortable speaking about them but may not have been the most knowledgeable among the coauthors of the paper(s) listed.
- (1) **Contributor** - I have little knowledge of VAMs beyond my contribution to the article(s) listed, for which I functioned mainly in a supporting role.

Which of the following disciplines do you consider to be your primary area of expertise?

- Economics** (including econometrics, finance, labor economics, public economics, industrial organization, and other similarly related fields)
- Education** (including content area instruction, classroom assessment, teacher preparation, educational leadership/administration, educational policy, and other similarly related fields)
- Quantitative Methodology** (including measurement, testing, psychometrics, statistics, analytical methods, and other similarly related fields)

On a 5-point scale, how would you rate your expertise in economics (including econometrics, finance, labor economics, public economics, industrial organization, and other similarly related

fields)? (consider your research and scholarship, your experience, your education and training, and the entire body of your professional work)

- (5) **Leading Expert**
- (4) **Expert**
- (3) **Experienced Scholar**
- (2) **Scholar**
- (1) **Early or Developing Scholar**
- (0) **N/A** (Economics is not my field)

On a 5-point scale, how would you rate your expertise in education (including content area instruction, classroom assessment, teacher preparation, educational leadership/administration, educational policy, and other similarly related fields)? (consider your research and scholarship, your experience, your education and training, and the entire body of your professional work)

- (5) **Leading Expert**
- (4) **Expert**
- (3) **Experienced Scholar**
- (2) **Scholar**
- (1) **Early or Developing Scholar**
- (0) **N/A** (Education is not my field)

On a 5-point scale, how would you rate your expertise in quantitative methodology (including measurement, testing, psychometrics, statistics, analytical methods, and other similarly related fields)? (consider your research and scholarship, your experience, your education and training, and the entire body of your professional work)

- (5) **Leading Expert**
- (4) **Expert**
- (3) **Experienced Scholar**
- (2) **Scholar**
- (1) **Early or Developing Scholar**
- (0) **N/A** (Methodology is not my field)

You will find some statements below which reflect some of the issues that have been discussed in the VAM literature. The statements have been organized into commonly discussed methodological and measurement concerns, and the sources of validity evidence described by the most recent edition of the joint *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014). For consistency and to minimize respondent error, the statements are positively phrased (i.e., in support of the use of value-added models) whenever possible.

Reliability & Precision

Definitions: **Reliability** is defined as the degree to which test- or measurement-based scores “are consistent over repeated applications of a measurement procedure [e.g., a VAM] and hence and inferred to be dependable and consistent” (AERA, APA, & NCME, p. 222-223). **Precision** is defined as the degree to which a score (e.g., a VAM score) is free from “the impact of measurement error on the outcome of the measurement” (AERA, APA, & NCME, p. 222)

Please indicate the degree to which you agree or disagree with the following statements, according to your best understanding of the scholarly evidence.

1) Value-added models (VAMs) produce estimates of teacher effectiveness that are reliable enough to be used for summative (e.g., outcomes-oriented) evaluation purposes.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

2) VAMs produce estimates of teacher effectiveness that are reliable enough to be used for summative evaluation purposes with high-stakes consequences attached (e.g., hiring, firing, merit pay, and other rewards and sanctions).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

3) VAM estimates are reliable enough to provide meaningful rankings of teachers (e.g., from highly effective to highly ineffective teachers).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

4) VAM estimates are precise enough to accurately identify teachers' levels of effectiveness in the tails of the effectiveness distribution (i.e., either highly effective or highly ineffective teachers).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

5) VAM estimates are precise enough to accurately identify teachers' levels of effectiveness in the middle of the effectiveness distribution (i.e., teachers that are neither highly effective nor highly ineffective).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

R&P) What other comments, if any, would you like to add regarding the reliability or precision of VAM estimates?

Bias & Construct Irrelevant Variance (CIV)

Definitions: **Construct irrelevant variance (CIV)** is defined as the variance in scores (e.g., VAM scores) “that is attributable to extraneous factors that distort the meaning of the scores and thereby decrease the validity of the proposed interpretation” (AERA, APA, & NCME, p. 217).

Bias, in this section, is defined as “systematic error in a test [e.g., VAM] score” (AERA, APA, & NCME, p. 216) **Please indicate the degree to which you agree or disagree with the following statements**, according to your best understanding of the scholarly evidence.

6) Value-added models (VAMs) produce unbiased estimates of teacher effectiveness independent of the ways teachers and students are assigned to classrooms.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

7) VAMs produce unbiased estimates of teacher effectiveness with statistical controls (e.g., including student attributes as covariates in the model).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

8) VAMs produce unbiased estimates of teacher effectiveness without statistical controls (e.g., models which include no covariates, using only prior achievement as controls).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

9) VAMs correctly identify the variance in student test scores that can be attributed to an individual teacher (i.e., estimates are not confounded by variance from other sources).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

10) VAMs produce accurate estimates of teacher effectiveness regardless of missing data.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

11) When missing data is present, VAMs produce accurate estimates of teacher effectiveness regardless of whether data are missing at random.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

B/CIV) What other comments, if any, would you like to add regarding CIV/bias or missing data, as related to VAM estimates?

Test Content (Content-Related Evidence of Validity)

Please indicate the degree to which you agree or disagree with the following statements, according to your best understanding of the scholarly evidence.

12) Value-added models (VAMs) produce accurate estimates of teacher effectiveness independent of test selection (i.e., using scores from one achievement test to generate VAM estimates as opposed to using a different test of the same content domain).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

13) The achievement tests currently used to generate VAM estimates are capable of measuring student growth in achievement over time.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

14) The achievement tests currently used to generate VAM estimates are capable of measuring a teacher's unique contribution to student growth in achievement over time.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

15) VAMs produce accurate estimates of teacher effectiveness when student prior achievement in one subject area (e.g., reading or mathematics) is used to estimate a teacher's value-added in a different subject area (e.g., science or social studies).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

16) VAMs produce accurate estimates of teacher effectiveness when prior achievement in a non-adjacent grade level is used to estimate a teacher's value-added (e.g., using a grade 4 science test to estimate a grade 8 science teacher's value-added).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

TC) What other comments, if any, would you like to add regarding the achievement tests used to generate VAM estimates?

Relationships to Other Variables (Concurrent- and Predictive-Related Evidence of Validity)

Please indicate the degree to which you agree or disagree with the following statements, according to your best understanding of the scholarly evidence.

17) Correlations observed between value-added model (VAM) estimates and long-term student outcomes (e.g., graduation rate, post-secondary education outcomes, and lifetime earnings) provide convincing evidence for the use of VAMs.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

18) Correlations between VAM estimates and supervisor observation scores provide convincing evidence for the use of VAMs.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

19) VAMs measure a substantively different aspect of teacher effectiveness than classroom observations, thus VAM estimates and supervisor observation scores do not need to correlate.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

20) Correlations between VAM estimates and student survey scores provide convincing evidence for the use of VAMs.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

21) VAMs measure a substantively different aspect of teacher effectiveness than student surveys, thus VAM estimates and student survey scores do not need to correlate.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

ROV) What other comments, if any, would you like to add regarding observed relationships between VAM estimates and these or other teacher effectiveness variables?

Intended Consequences

Please indicate the degree to which you agree or disagree with the following statements, according to your best understanding of the scholarly evidence.

22) The use of value-added models (VAMs) encourages teachers to increase their professional efforts in the classroom.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

23) States and districts which have relied on VAMs to determine personnel decisions (e.g., hiring, firing, merit pay, and other rewards and sanctions) have seen instructional quality increase.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

24) States and districts which have relied on VAMs to determine personnel decisions (e.g., hiring, firing, merit pay, and other rewards and sanctions) have seen student achievement increase.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

25) The use of VAMs encourages the most effective teachers to teach the students with the greatest educational needs.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

26) VAMs promote transparency in the teacher evaluation process.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

IC) What other comments, if any, would you like to add regarding these or any other intended consequences of VAM use?

Unintended Consequences

Please indicate the degree to which you agree or disagree with the following statements, according to your best understanding of the scholarly evidence. Note: Unlike previous pages, most of the items on this page are phrased *negatively* and reverse coded (such that “Strongly agree” = 1).

27) Teachers are less inclined to teach certain subject areas (e.g., mathematics or reading) for which teachers may be held accountable using value-added models (VAMs).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

28) Teachers are less inclined to teach certain grade levels (e.g., grade 3 through grade 8) for which teachers may be held accountable using VAMs.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

29) Teachers are less inclined to teach students with histories of low academic performance so that such students cannot potentially suppress teachers' VAM estimates.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

30) Teachers are less inclined to teach students with histories of high academic performance so that such students cannot potentially suppress teachers' VAM estimates (e.g., through potential ceiling effects).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

31) Teachers are less inclined to work with other teachers in teams or collaborative groups as a result of VAM use.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

32) The use of VAMs has led to increased teacher competition.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

33) The use of VAMs has decreased teacher morale.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

34) Teachers are inclined to switch schools, districts, or states to work in settings where VAM estimates might not matter as much (e.g., for hiring, firing, merit pay, and other rewards and sanctions).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

35) Teachers are inclined to leave the teaching profession to avoid the perceived or real threats of consequences attached to VAMs.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

UC) What other comments would you like to add regarding these or any other unintended consequences of VAM use?

VAM Use

Please indicate the degree to which you agree or disagree with the following statements, according to your best understanding of the scholarly evidence.

36) The benefits and intended consequences of using value-added models (VAMs) outweigh any drawbacks or unintended consequences of their use.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

- 37) VAMs should continue to be used to evaluate teachers.
- Strongly agree
 - Somewhat agree
 - Neither agree nor disagree
 - Somewhat disagree
 - Strongly disagree
 - Don't know / Uncertain
- 38) VAMs can and should be used to evaluate teachers of every grade level.
- Strongly agree
 - Somewhat agree
 - Neither agree nor disagree
 - Somewhat disagree
 - Strongly disagree
 - Don't know / Uncertain
- 39) VAMs can and should be used to evaluate teachers of every content area.
- Strongly agree
 - Somewhat agree
 - Neither agree nor disagree
 - Somewhat disagree
 - Strongly disagree
 - Don't know / Uncertain
- 40) VAMs are superior to other available methods of evaluating teacher effectiveness.
- Strongly agree
 - Somewhat agree
 - Neither agree nor disagree
 - Somewhat disagree
 - Strongly disagree
 - Don't know / Uncertain
- 41) VAMs should represent the most important component of teacher evaluation systems (e.g., most heavily weighted).
- Strongly agree
 - Somewhat agree
 - Neither agree nor disagree
 - Somewhat disagree
 - Strongly disagree
 - Don't know / Uncertain

42) VAMs can and should be used to help teachers improve instruction.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

43) VAMs should continue to be used to make professional development decisions for teachers (e.g., to determine which teachers need what types of professional development).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

44) VAMs should continue to be used to hold teachers accountable for student learning with consequences attached to teachers' VAM estimates (e.g., merit pay, teacher tenure, termination, and other rewards and sanctions).

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- Don't know / Uncertain

VU) What other comments (if any) would you like to add regarding these or any other uses of VAMs?

Free-Response Items

Please respond to the following free-response items. You may provide as much or as little detail as you like.

45) Given that the Every Student Succeeds Act (ESSA) removes the specific requirements which have led to the widespread use of value-added models (VAMs), and provides flexibility to states and districts to modify their current teacher evaluation methods, what advice would you give an interested state or district regarding their VAM adoption and use?

46) If you had the opportunity to make a brief comment or recommendation about VAMs or VAM use to the current US Secretary of Education, Betsy DeVos, what would you say?

Thank you for your participation!