



MISSISSIPPI DEPARTMENT OF EDUCATION

Carey M. Wright, Ed.D.
State Superintendent of Education

July 10, 2015

James Barber, Executive Director
Joint Legislative PEER Committee
Woolfolk Building
501 North West Street, Suite 301A
Jackson, Mississippi 39201

Dear Mr. Barber,

The Mississippi Department of Education (MDE) has reviewed the PEER Committee report "A Review of Accountability Standards of the Mississippi Department of Education." While the evaluation is well-written and organized, the evaluators' chain-of-logic and evidence does not support many of the conclusions. The report does not provide either empirical or research-based evidence to support many conclusions. Further, technical misinterpretations suggest subject-matter experts in the fields of psychometrics, accountability systems, and statistics may not have fully reviewed the findings or recommendations.

Specifically, the report outlined the following concerns and public perceptions related to the performance standards, which are part of the Mississippi Statewide Accountability System. MDE's rationale regarding the conclusions follows each of these concerns.

PEER Concern: MDE's calculation of students' scores on assessments, upon which the accountability standards are based, does not offer a clear picture of how schools and districts are actually performing, nor does it show whether schools and districts have achieved what they are supposed to have achieved.

MDE Response: Many state education accountability systems use performance categories precisely because it clearly incorporates the assessment standard into the system. Indeed, Mississippi's criterion referenced tests are designed for the explicit purpose of classifying with the highest degree of precision student performance with respect to these standards. Any assertion that scale scores should be the primary driver of accountability determinations is incomplete without consideration of the error associated with interpreting unit level changes in scores which may be within the standard error of measurement. In other words, scale score based indicators are prone to false-precision that is not recognized in the recommendations.

PEER Concern: The U. S. Department of Education's voluntary waiver that allows schools to choose the higher of two years' grades (for the 2012-2013 and 2013-2014 school years) does not make it possible to compare a school's accountability grade from one year to another or to compare the grades of several schools or districts over time.

MDE Response: MDE has provided data on its public reporting website with all the resulting calculations, with and without the "waived" result, since the State Board of Education (SBE) approved the results on October 17, 2014. The file is available on MDE's public reporting webpage, as well as through MDE's official [news release](#) regarding 2014 Accountability Results.

PEER Concern: Certain multipliers included in a school's or district's accountability grade calculation result in inflation of schools' and districts' grades.

MDE Response: The growth metric includes weighting achievement levels in order to incentivize improvement in student performance from one year to the next. The growth indicators are combined with status indicators to create an overall score. This compensatory design reduces the misclassification of schools and districts. The evaluators' recommendation would increase the misclassification of the accountability determinations, thus resulting in spurious results. The evaluators' provided no empirical data supporting their recommendation. Further, they provided neither actual nor simulated analytics demonstrating that additional conjunctive decisions would have misclassification (i.e., Kappa) indices similar the SBE's current design.

PEER Concern: Some of MDE's Accountability Task Force's decision-making practices do not align with the overall purpose of the accountability standards.

MDE Response: The agency selected and applied a standard-setting approach that has extensive research supporting the validity of the approach. The agency used panelists who were subject matter experts in the area of educational leadership and management (e.g., principals, supervisors, superintendents) to recommend cut-scores for the State Board of Education's consideration.

PEER Concern: MDE's ability to create assessment benchmarks and cut-points or the accountability grades gives the department an inappropriate amount of control over these grades.

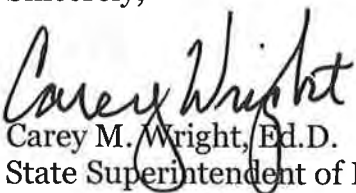
MDE Response: The literature has demonstrated the volatility of growth-based metrics. This volatility is in part ameliorated by including status indicators, using multiple data waves, and ensuring only reliable assessment results (low standard error of measure) are included. The inclusion of growth-based metrics with status-

based metrics increases the robustness of the accountability design. The addition of a growth-metric improves the goodness-of-fit of the accountability model to predict accountability ratings.

The documents that follow support the existing accountability system. Each recommendation from the PEER Committee report is addressed. The resumes for MDE's Technical Advisory Committee (TAC) and key members of the Technical Review Team provide evidence of the quality of the team vetting the system. It should be noted that these individuals have extensive experience in designing accountability and assessment systems and are recognized as national experts in their fields. I requested the resumes of those who conducted the review and only received the degrees they obtained, as shown in the attachment. I question the technical sufficiency of the PEER Review Team, which is needed to evaluate, draw conclusions, and formulate design recommendations for state assessment and accountability systems.

The performance standards of the Mississippi Statewide Accountability System are a high-quality set of indicators of school effectiveness, intended to improve student outcomes. The system, developed through a multi-year deliberate process of stakeholder feedback and research-based best practice, reflects a very different philosophical approach to school accountability than the report illustrates. The MDE has worked diligently to ensure reliability and validity in these indicators of school success, incorporating both state and federal statute into one coherent model to measure school performance.

Sincerely,



Carey M. Wright, Ed.D.
State Superintendent of Education

Attachments: MDE Response to PEER Accountability Report
Vitas/Resumes of:

- Joint Legislative PEER Committee
- MDE Technical Advisory Committee for Assessment and Accountability
- MDE Technical Review Team, External Assessment and Accountability Members

Mississippi Department of Education
Response to Recommendations from the PEER Committee report “A
Review of Accountability Standards of the Mississippi Department of
Education”

Recommendation 1. In order for a school's or district's student proficiency to be represented accurately by its accountability grade, MDE should report performance grades that reflect student assessment score data as closely as possible. This could be done by:

- eliminating the use of the four achievement categories (minimal, basic, proficient, and advanced); or,
- reporting scale scores divided by total possible scale score points (in the form of a percentage).

To accomplish this, the Legislature should amend MISS. CODE ANN. 37-17-6(5)(c)(i) (Note: When *proficiency* is referenced in other recommendations in this report, it is with the assumption that an accurate proficiency measure will be utilized.)

MDE Response: The report raises objections about the use of performance categories to signal student performance, indicating that this is less preferable to an indicator based solely on scale scores (pp. 29-30). However, it is important to note that many state education accountability systems throughout the nation use performance categories precisely because it clearly incorporates the assessment standard into the system. Indeed, Mississippi's criterion referenced tests are designed for the explicit purpose of classifying with the highest degree of precision student performance with respect to these standards. Furthermore, any assertion that scale scores should be the primary driver of accountability determinations is incomplete without consideration of the error associated with interpreting unit level changes in scores which may be within the standard error of measurement. In other words, **scale score based indicators are prone to false-precision that is not recognized in the recommendations.**

Recommendation 2. In order to communicate and report student proficiency and student growth accurately and to prevent either proficiency or growth from greatly affecting a school's or district's accountability grade, MDE should separate proficiency and growth into two separate grades.

MDE could do so by assigning a letter grade (A thru F) for proficiency, followed by another indicator to represent growth. The department could use a letter grade to demonstrate proficiency and an arrow that indicates direction to reflect whether a school has made adequate learning gains. For example, a school that made learning gains and earned a B in proficiency would have a grade of B+. However, a school that earned a B in proficiency, but did not make adequate learning gains, would have a grade of B-.

For the separation of scores to take place, the Legislature would need to amend MISS. CODE ANN. Section 37-17-6 (g) (1972) to allow for separate proficiency and growth indicators.

MDE Response: A central assertion of the report authors is that the influence of growth corrupts or clouds the accountability outcomes. For example, on page 29 the authors state, “Mississippi’s standardized tests are carefully constructed to ensure that a student has mastered a certain level of competency; those tests alone provide the criterion/ standard that should measure school performance.” Again on page 33 the authors reason, “If the purpose of the accountability standards is to improve student achievement and increase the level of accountability of schools and districts, then more emphasis should be placed on proficiency.” Taken to its natural conclusion, the application of this statement creates a system that systematically disadvantages schools, especially those that serve traditionally low achieving students. A well-designed and fair system must include multiple indicators including, especially, elements that reward student growth. Furthermore, to the extent that equity (or improving performance of at-risk students) is a priority outcome, indicators must be included that signal such progress and reward ongoing efforts. To this point, the Council of Chief State School Officers (CCSSO) document entitled *CCSSO Roadmap for Next-Generation State Accountability Principles*, which was developed by a representative group of state education leaders and experts in the field of assessment and accountability, explicitly promotes as principle 3, “Focus on student outcomes on a variety of indicators including those of both status and growth” (see: http://www.ccsso.org/Documents/2011/Roadmap_for_Next-Generation_Accountability_2011.pdf). From the same seminal report, the authors identify the following option for states, “Focus particularly on and weigh more heavily the achievement of the lowest-performing students.” In short, there is a strong rationale for and extensive practice nationally in support of design decisions represented in the Mississippi system. Conversely, there is a growing recognition nationally, that systems that rely exclusively or heavily on status test performance do not represent best practice.

Recommendation 3. To ensure that a school's or district's growth is represented accurately in its accountability grade, MDE should indicate growth by a student's improvement from one year to the next in the accurate proficiency grade. MDE uses growth multipliers of 1, 1.2, or 1.25 to indicate greater growth, but any multiplier or incentive that alters an original score takes a rating farther away from accurately demonstrating true performance.

MDE Response: On page 34 of the report, there are multiple references to growth being counted three times in the model. Growth is only counted twice, and then ONLY if the student is in the Low subgroup. The support for the three-time argument is that a student who performs at proficient or advanced, by nature, is growing. However, that is not always the case. In fact, statewide results for school year 2013-14, of the students meeting proficiency, approximately 80% of students met both proficiency and growth, 11% met proficiency but not growth, and 9% met proficiency and exceeded growth.

On page 37, in the last sentence of the last paragraph, the following statement is made: “if those incentives obfuscate data regarding actual student performance, the ultimate goal of improving student achievement and increasing school and district accountability

has not been reached.” Reviewing statewide results from school year 2013-14 reveals that, if the weights are removed from the growth calculations, approximately one point makes the difference in weighted and unweighted values. One point of increase is more incentive than obfuscation. The report also asserts that “MDE could provide no rationale for the values of the weights”; however, incentivizing growth is the rationale, as noted in the report.

On page 36 the authors **suggest two approaches for “presenting growth.”** The first is to report proficiency rates annually and focus on changes in rates. Simply put, this is not growth. This refers to a trend analysis which is not based on matched cohorts of the same students. Achievement differences over time using such a method cannot be disentangled from annual changes in school enrollment. The authors further suggest an approach whereby the scale score is reported as a percentage of the total possible scale score. This is not technically appropriate or defensible. Such a practice completely departs from the measurement model used for the tests—Item Response Theory (IRT)—and would substantially distort precision and meaningful score interpretation.

Recommendation 4. To ensure that a school's or district's grade for a given year is a direct representation of that school's or district's performance for that year, MDE should instruct schools and districts to report and publicize not only their official grade, but **also their “paused” or “waived” grades in any school year that is** considered a transitional year. Allowing schools and district the opportunity to publicize the better grade of two years, or an outdated school grade, does not provide a clear picture of current performance.

Further, to ensure that schools' and districts' grades can be reliably compared to those of other schools or districts for that year and that a single school or district can analyze its performance over a period of time, MDE should report schools' and districts' grades using the same accountability standards (as opposed to a previous year's standards or a previous year's grades).

MDE Response: The report raises concerns about being able to compare scores from year to year (pp. 39-40). However, MDE has released to the public all the SY 13-14 calculations, with and without the “pause” or “waived” grade. Further, MDE intends to run the SY14-15 model, as the values are required by the U.S. Department of Education for federal accountability purposes. The first sentence of the last full paragraph on page **40 indicates that “the public might not have actually seen the accountability results as MDE has purported they are calculated.”** MDE has provided data on its public reporting website with all the resulting calculations, **with and without the “waived”** result, since the State Board of Education approved the results on October 17, 2014. The file is available on MDE public reporting webpage, as well as through MDE’s **official news release** regarding 2014 Accountability Results.

Recommendation 5. To ensure that the A through F cut-points and assessment benchmarks are directly related to student mastery over material, MDE should develop a defensible criterion for being “proficient.”

MDE Response: The method for setting standards for both the state tests and the state accountability system **seem to be questioned as too ‘subjective’ rather than ‘criterion-based’ (p.45). The process for setting standards for the state tests follows** established, research-based protocols used broadly for state achievement tests. The Bookmarking (Lewis, Mitzel, & Green, 1996) and the Contrasting Group (Livingston & Zieky, 1982) methods used to establish the academic achievement standards (AAS) are internationally recognized approaches. These methods were used in a majority of statewide assessment systems reviewed **by the United States Department of Education’s** Peer Review process in meeting the requirements of 34 C.F.R. 200. **“The performance standards provide qualitative descriptions of the intended distinctions between adjacent levels of performance. The cut scores are points on the score scale, with one cut score associated with each performance standard (Cizek, 2001, pg. 55). This process relies on both judgments and data. It is precisely because of this process, that a criterion-based claim can be supported for the assessments. Moreover, utilizing a broad-based task force to make judgments based in part on data, as is the practice in Mississippi, to reach decisions about the accountability system standards is a common and acceptable practice nationally to identify performance criteria. If the authors have another proposal for establishing the key performance criteria for either the tests or the accountability system, it was not apparent to the reader. The process evidence demonstrates the technical quality of the standard-setting methodology and the nationally recognized experts associated with implementing the procedures. Claims asserting the “defendable criterion” are neither supported by psychometric experts or empirical evidence.**

Recommendation 6. To ensure that the accountability standards accomplish what they are designed to accomplish, MDE should ensure that task force recommendations support the purpose of the accountability standards so that appropriate changes, where necessary, can be made.

MDE Response: The practice of engaging stakeholders in the decision-making process is one that is widely used across various disciplines. Allowing a diverse group of current practitioners to have a voice in the development of the accountability system helps to strengthen its efficiency, reduce risks, and increases transparency and understanding around its design. The role of the Accountability Task Force (ATF) is to advise MDE, State Commission on School Accreditation, and State Board of Education (SBE) on the business rules and practices that improve the implementation of the accountability system and its measure of district and school performance. The ATF reviews and discusses thoroughly the purpose of the accountability standards as well as data, where necessary, to ensure the decisions are equitable for all districts and schools across the state. As noted on page 46 of the report, all of the recommendations generated through the work of the ATF are vetted through the Technical Review Team and Technical Advisory Committee to determine the soundness of the decisions being brought forward as well as to ensure the decisions are statistically/technically defensible. In evidence of the quality of the vetting process, the resumes for the Technical Advisory Committee, as well as for the members of the Technical Review Team with expertise in national educational accountability research, are provided at the end of this response. Prior to the approval of these recommendations by the SBE, the general public are afforded an opportunity to provide feedback to the SBE on concerns of the proposed

recommendations through the Administrative Procedures Act (APA) process. This feedback are considered and changes are made as necessary prior to the final ruling of the SBE.

Recommendation 7. In the best interest of the students and to acknowledge the distinct honor of successfully completing high school, MDE should develop a method to ensure that the changes made to the graduation options are equivalent and comparable to a standard/regular high school diploma.

MDE Response: The agency applied weighting to both academic achievement results and graduation rates consistent with SBE's policy. The emphasis on students graduating on time with a regular diploma is enumerated in the weighting. The emphasis on students graduating from high school and either entering the workforce or attending college is supported by the research. An extensive body of research exists showing that lifetime earnings of high school graduates are significantly higher than for those that dropout of schools (Card, 1999; Cheeseman, & Newburger, E., 2014). Further, the research also shows strong evidence that high school dropouts are incarcerated (Anderson, 1999), unemployed, and use government support services (Grossman, 2006) at rates much higher than graduates. The graduation options that result in the issuance of a standard, high school diploma use multiple indicators (i.e., course credit earned, standardized assessment results). Other states' policies allow for additional performance evidence to be considered, such as remediation opportunities, attendance, grade-point-average, portfolios, electives, or occupational readiness examinations, as well as other nationally recognized assessments (see Education Commission of the States, Indiana Department of Education). These nationwide changes reflect the divergence of the K-12 systems in preparing students to be ready to enter the workforce or to move to post-secondary education. The high-stakes "gate keeper" test policies of the late 1990s as the only indicator of post-secondary readiness have been determined to be fraught with negative consequences (e.g., disadvantaging students with disabilities and English language learners). Further, the MDE ensures that changes made to graduation options are equivalent and comparable to a standard/regular high school diploma through the input of our Technical Advisory Committee (resumes attached) and the use of national research in the field, as well as the thorough vetting of decisions through the Commission on School Accreditation, the State Board of Education, and public comments garnered through Administrative Procedures Act (APA) process.

Recommendation 8. The Legislature should enact legislation requiring that the Mississippi Department of Education submit any proposed changes to the school accountability standards one year before those standards may become effective. Such submission should also include a statement of estimated economic impact detailing how the proposed changes could impact the development of recommendations for the funding of the adequate education program. This is important because school districts' accountability grades are figured into the MAEP formula and any changes in the way that a "successful" district (currently, a district receiving a C accountability grade) is defined will affect the calculation of the MAEP funding formula and thereby affect the

amount of funding requested by MDE and ultimately the amount of funding received by school districts.

MDE Response: As authorized through section 203(2) of the Mississippi Constitution, the State Board of Education formulates “policies according to the law for implementation by the State Department of Education.” Following statewide procedure, the SBE finalizes policy through the Administrative Procedures Act (APA) process, as codified in MISS. CODE ANN. § 25-43-1 through 4. The legislation governing this process provides for public comment and review of any changes and potential impacts.

Overall, MDE is disappointed in the quality of the report. Based on the extensive and research-based approach used to develop the Mississippi Statewide Accountability System, MDE has tremendous confidence it accurately reflects the progress of districts and schools in a most transparent manner.

Attachments: Vitas/Resumes

Joint Legislative PEER Committee

Education Accountability Team (provided by PEER to MDE)

MDE Technical Advisory Committee for Assessment and Accountability

Dr. Chad Buckendahl

Dr. Chris Domaleski, Chair

Dr. George Englehard

Dr. Gerunda Hughes

Mr. Robert Lee

**MDE Technical Review Team, External Assessment and Accountability
Members**

Dr. JP Beaudoin

Dr. Christy Hovanetz

**Joint Legislative PEER Committee
Staff Biographies**

Education Accountability Project

Project Team

David Pray

Principal Analyst
Certified Public Accountant
Master of Business Administration, Mississippi College
Bachelor of Business Administration, University of Mississippi

Jenell Ward

Analyst
Master of Public Health, University of Southern Mississippi
Bachelor of Science in Nutrition and Dietetics, University of Southern Mississippi

Julie Winkeljohn

Senior Analyst
Master of Theological Studies, Spring Hill College
Master of Business Administration, Mississippi State University
Bachelor of Arts in Communication, Mississippi State University

PEER Staff Providing Administrative, Quality Assurance, and Editing Support

Max K. Arinder

Executive Director
Doctor of Philosophy in Experimental Psychology, University of Southern Mississippi
Master of Arts in Psychology, University of Southern Mississippi
Bachelor of Arts in Psychology, University of Southern Mississippi

James A. Barber

Deputy Director
Master of Business Administration, University of Southern Mississippi
Bachelor of Business Administration, University of Southern Mississippi

Jennifer Sebren

Principal Analyst
Master of Science in Industrial/Organizational Psychology, University of Tennessee at Chattanooga
Bachelor of Science in Psychology, Mississippi College

Ava Welborn

Editing Coordinator/Quality Assurance Manager
Master of Science in English, University of Southern Mississippi
Bachelor of Arts in English, University of Southern Mississippi

Chad W. Buckendahl

2467 Cordoba Bluff Ct
Las Vegas, NV 89135
Phone/Fax: (702) 586-7386
Cell phone: (402) 770-0085
Email: drcbuck@gmail.com

Education

- Ph.D. 2000, Quantitative and Qualitative Methods in Education,
University of Nebraska, Lincoln
- M.L.S. 1996, Legal Studies, University of Nebraska College of Law
- B.A. 1994, Political Science, University of Nebraska, Lincoln

Professional Experience

- 2007 – pres. Senior Psychometrician and Director of Strategic Partnerships,
Alpine Testing Solutions.
Responsibilities: Develop and maintain client relationships,
provide psychometric consultation, manage applied research
projects, and conduct applied research for testing programs in
education, licensure, and certification.
- 1998 – 2007 Director (2002-2007), Assistant Director (2000-2002), Research
Associate (1998-2000), Buros Institute for Assessment
Consultation and Outreach, University of Nebraska, Lincoln.
Responsibilities: Sought external funding, directed applied
research and provided measurement consultation for proprietary,
non-commercial testing programs in education, licensure,
certification, and admissions.
- 1990 – 1997 Computer Programmer (1994-1997), Survey Researcher (1990-
1994), The Gallup Organization, Inc.
Responsibilities: Collected survey data for a variety of clients,
authored training manuals, conducted research on in-person survey
participation rates, and programmed surveys in 'C' for CATI and
CAPI data collection.

Memberships and Professional Service

American Educational Research Association (AERA)

Member, 1997 to 2010, Division D, Professional Licensure and Certification SIG

Reviewer, 1999 to 2010, Division D, Professional Licensure and Certification SIG

2006 Co-winner of Division D Award for Significant Contribution to Educational Measurement and Research Methodology

Institute for Credentialing Excellence (ICE)

Member, 2009 to present

NCCA Commissioner (Psychometric Reviewer), 2010 to present

NCCA, Chair, 2013 to present, Co-chair, 2011 to 2013

International Test Commission (ITC)

Member, 2002 to present

National Council on Measurement in Education (NCME)

Member, 1997 to present

Reviewer, 1999 to present

Co-chair of membership committee, 2004-05, member 2004-07

Co-chair of program committee, 2005-06

Chair of outreach committee, 2007-08, member 2007-10

Northern Rocky Mountain Educational Research Association (NRMERA)

Member, 2002 to 2012

Executive Board Member, 2002-05, Program Chair, 2003, President-Elect, 2005-06, President, 2006-07, Past President, 2007-08

Phi Delta Kappa (PDK)

Member, 2000 to 2010

Journal Editor/Reviewer

Editor: *Journal of Applied Testing Technology*, 2008-2015

Associate Editor: *Applied Measurement in Education*, 2006-2010

Guest Editor: *Applied Measurement in Education* (Summer 2005)

Editorial Board: *Applied Measurement in Education* 2005-06, *Third Education Group Review*, 2004-05, *International Journal of Testing*, 2014-present

Occasional reviewer: *Applied Measurement in Education*; *Applied Psychological Measurement*; *Assessment in Education: Principles, Policies, and Practice*; *Educational Assessment*; *Educational Measurement: Issues and Practice*; *Environmental Health Perspectives*; *International Journal of Testing*; *Journal of the American Dental Association*; *Language Assessment Quarterly*; *Measurement: Interdisciplinary Research and Perspectives*; *Mid-Western Educational Researcher*

Publications

Journal Articles

1. Davis-Becker, S. L. & Buckendahl, C. W. (2013). A proposed framework for evaluating alignment studies. *Educational Measurement: Issues and Practice*, 31(1), 23-33.
2. Ferdous, A. A. & Buckendahl, C. W. (2013). Evaluating panelists' standard setting perceptions in a developing nation. *International Journal of Testing*, 13(1), 4-18.
3. Davis-Becker, S. L. & Buckendahl, C. W. (2013). Identifying and evaluating external validity evidence for passing scores. *International Journal of Testing*, 13(1), 50-64.
4. Lim, G. S., Geranpayeh, A., Khalifa, H., & Buckendahl, C. W. (2013). Standard setting to an international reference framework: Implications for theory and practice. *International Journal of Testing*, 13(1), 32-49.
5. Buckendahl, C. W. & Davis-Becker, S. L. (2012). The appropriateness and use of domain critical errors. *Practical Assessment, Research, & Evaluation*, 17(13), 1-12.
6. Davis-Becker, S., Buckendahl, C., & Gerrow, J. (2011). Implications of random ordering on the bookmark standard setting method. *International Journal of Testing*, 11(1), 24-37.
7. Buckendahl, C., Ferdous, A. & Gerrow, J. (2010). Recommending cut scores with a subset of items: An empirical illustration. *Practical Assessment, Research & Evaluation*, 15(6). Available online: <http://pareonline.net/getvn.asp?v=15&n=6>.
8. Davis, S. L. & Buckendahl, C. W. (2009). An overview of the 2004-2008 NAEP evaluation: Design and research studies. *Applied Measurement in Education*, 22(4), 317-320.
9. Buckendahl, C. W., Plake, B. S., & Davis, S. L. (2009). Conducting a lifecycle audit of the National Assessment of Educational Progress. *Applied Measurement in Education*, 22(4), 321-338.
10. Davis, S., Buckendahl, C. & Plake, B. (2008). When adaptation is not an option: An application of multilingual standard setting. *Journal of Educational Measurement*, 45(3), 287-304.

11. Norman, R. & Buckendahl, C. (2008). Determining sufficient measurement opportunities when using multiple cut scores. *Educational Measurement: Issues and Practice*, 27(1), 37-46.
12. O'Neill, T., Buckendahl, C., Taylor, L., & Plake, B. (2007). Recommending a nursing-specific passing standard for the IELTS examination. *Language Assessment Quarterly*, 4(4), 295-317.
13. Hammond, D. & Buckendahl, C. (2006). Measurement trends in clinical licensure testing programs. *Journal of the American Dental Association*, 137, 30-42.
14. Buckendahl, C. (2005). Practical and perceived challenges to instructionally supportive accountability tests. *Measurement: Interdisciplinary Research and Perspectives*, 3(3), 188-201.
15. Yang, Y., Buckendahl, C., Juszkiwicz, P., & Bhola, D. (2005). Evaluating computer automated scoring: Issues, methods, and an empirical illustration. *Journal of Applied Testing Technology* (http://www.testpublishers.org/atp_journal.htm).
16. Buckendahl, C., Huynh, H., Siskind, T., & Saunders, J. (2005). A case study of vertically-moderated standard setting for a state science assessment program. *Applied Measurement in Education*, 18(1), 83-98.
17. Boss, T., Endorf, D., & Buckendahl, C. (2005). Developing a local assessment system for state accountability: A district's reflections. *The Researcher*, 19(1), 32-41.
18. Buckendahl, C., Plake, B. & Impara, J. (2004). A strategy for evaluating district developed assessments for state accountability. *Educational Measurement: Issues and Practices*, 23(2), 17-25.
19. Plake, B., Impara, J., & Buckendahl, C. (2004). Evaluating technical quality for district developed assessments. *Educational Measurement: Issues and Practices*, 23(2), 12-16.
20. Bhola, D., Impara, J., & Buckendahl, C. (2003). Aligning tests with states' content standards: Methods and issues. *Educational Measurement: Issues and Practices*, 22(3), 21-29.
21. Buckendahl, C., Impara, J., & Plake, B. (2002). District accountability without a state assessment: A proposed model. *Educational Measurement: Issues and Practices*, 21(4), 6-16.

22. Yang, Y., Buckendahl, C., Juskiewicz, P., & Bhola, D. (2002). A review of strategies for validating computer automated scoring. *Applied Measurement in Education, 15*(4), 391-412.
23. Buckendahl, C., Smith, R., Impara, J. & Plake, B. (2002). A comparison of Angoff and Bookmark standard setting methods. *Journal of Educational Measurement 39*(4), 253-263.
24. Arcot Sadanand, N., Samal, A., Soh, L., & Buckendahl, C. (under review). An adaptive placement test for introductory computer science courses. *Assessment and Evaluation in High Education*.
25. Buckendahl, C., Impara, J., Giraud, G., & Irwin, P. (2000). The consequences of judges making advanced estimates of impact on a cut score. (ERIC Clearinghouse on Assessment and Evaluation TM031248).
26. Rosson, B., Buckendahl, C., & Impara, J. (2000). Professional Activities and Knowledge of the Building Architectural Engineer. *Journal of Architectural Engineering*.
27. Giraud, G., Impara, J., & Buckendahl, C. (2000). Making the cut in public schools: Alternative methods for standard setting. *Educational Assessment, 6*(4), 291-304.
28. Buckendahl, C., Plake, B., & Impara, J. (1999). Setting minimum passing scores on high-stakes assessments that combine selected and constructed response formats. (ERIC Clearinghouse on Assessment and Evaluation TM029986).
29. Thompson, R., Christiensen, E., Jackson, S., Wyatt, J., Colman, R., Peterson, R., Wilcox, B., & Buckendahl, C. (1999). Parent attitudes and discipline practices: Profiles and correlates in a nationally representative sample. *Journal of Child Development, 4*(4), 316-330.
30. Jackson, S., Thompson, R., Christiansen, E., Colman, R., Wyatt, J., Buckendahl, C., Wilcox, B., & Peterson, R. (1999). Predicting abuse-prone parental attitudes and discipline practices in a nationally representative sample. *Child Abuse & Neglect, 23*(1), 15-29.

Books, Book Chapters or Monographs

1. Buckendahl, C. W. (in preparation). Understanding credentialing's role in workforce development. In S. Davis-Becker & C. Buckendahl (Eds.), *Testing in the professions*. New York, NY: Routledge.

2. Buckendahl, C. W. (in preparation). Public perceptions of educational assessment: Consistently positive? In G. Brown & L. Irvin (Eds.), *Parents' perceptions of assessment*. Melbourne, Australia: Taylor & Francis.
3. Foley, B. P. & Buckendahl, C. W. (2013). Using visual displays to inform assessment design and development. In G. Schraw, M. McCrudden, & D. Robinson (eds.), *Learning through visual displays*. Charlotte, NC: Information Age Publishing.
4. Buckendahl, C. & Davis-Becker, S. (2012). Setting passing standards for credentialing examinations. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 485-502). New York, NY: Routledge.
5. Davis-Becker, S. & Buckendahl C. (2011). Integration of cognitive demand into licensure and certification exam development. In G. Schraw (Ed.), *Assessment of higher order thinking skills* (pp. 303-326). Charlotte, NC: Information Age.
6. Buckendahl, C. & Foley, B. (2011). A high-stakes uses of intelligence testing: A forensic case study. In J. Bovaird, K. Geisinger, & C. Buckendahl (Eds.), *High stakes testing in education: Science and practice in K-12 setting* (pp. 191-210). Washington, DC: American Psychological Association.
7. Ferdous, A., Bechard, S., & Buckendahl, C. (2011). Setting performance standards on alternate assessments for students with disabilities. In J. Bovaird, K. Geisinger, & C. Buckendahl (Eds.), *High stakes testing in education: Science and practice in K-12 setting* (pp. 43-58). Washington, DC: American Psychological Association.
8. Buckendahl, C., Davis, S., Plake, B., Sireci, S., Hambleton, R., Zenisky, A. & Wells, C. (2009). *Evaluation of the National Assessment of Educational Progress: Final Report*. Washington, DC: U.S. Department of Education
9. Buckendahl, C. & Plake, B. (2006). Evaluating Tests. In S. Downing & T. Haladyna (Eds.), *Handbook of Test Development* (pp.725-738). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
10. Buckendahl, C. & Hunt, R. (2005). Whose Rules? The relation between the “rules” and “laws” of testing. In R. P. Phelps (Ed.), *Defending Standardized Testing* (pp.147-158). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Presentations

International and National

1. Buckendahl, C. W. & Foley, B. P. (April, 2015). *Policy linking as cut score moderation: Considerations for practice*. Paper presented at the annual meeting of the National Council on Measurement in Education. Chicago, IL
2. Buckendahl, C. W. (April, 2015). *Accreditation of credentialing programs: An analogous model for educational assessment?* Paper presented at the annual meeting of the National Council on Measurement in Education. Chicago, IL.
3. Michaels, H., Geddes, M., & Buckendahl, C. (March, 2015). Evaluating standard setting studies. Presentation at the annual meeting of the Association of Test Publishers. Rancho Mirage, CA.
4. Buckendahl, C., Smith R., & Gerrow, J. (July, 2014). *Evaluating the impact of an intentional item pool release*. Paper presented at the biannual meeting of the International Test Commission. San Sebastian, Spain.
5. Buckendahl, C., Zenisky, A., van den Heuvel, J., & Davis-Becker, S. (July, 2014). *International perspectives on score reporting in credentialing: Best practices for providing meaningful feedback to examinees*. Paper presented at the biannual meeting of the International Test Commission. San Sebastian, Spain.
6. Lim, G. & Buckendahl, C. (July, 2014). *Multiple standard setting study outcomes: Empirical exercises informing theory*. Paper presented at the biannual meeting of the International Test Commission. San Sebastian, Spain.
7. Buckendahl, C., Sharp, C., & Parton, C. (June, 2014). College-ready, career-ready, and graduation eligibility: Challenges for EOC assessments. Presentation at the National Conference on Student Assessment. New Orleans, LA.
8. Buckendahl, C., Misner, R., Munson, L, & Rupp, S. (March, 2014). Singing for your supper: Are performance-based tests worth the price of admission. Presentation at the annual meeting of the Association of Test Publishers. Scottsdale, AZ.
9. Buckendahl, C. (April, 2013). *A practitioner's guide to validation framework development*. Paper presented at the annual meeting of the National Council on Measurement in Education. San Francisco, CA.
10. Bontempo, B., Buckendahl, C., Muckle, T., & Neustal, S. (February, 2013). Benchmarks: A (theoretically) comprehensive overview of industry standards and best practice guidelines. Presentation at the Association of Test Publishers' Innovations in Testing Conference. Ft. Lauderdale, FL.
11. Buckendahl, C. & Smith, R. (July, 2012). *Evaluating the release of an item bank to respond to security concerns*. Paper presented at the International Test Commission conference. Amsterdam, The Netherlands.

12. Davis-Becker, S. & Buckendahl, C. (July, 2012). *Enhancing test development through digital technology: Blended and distributed modes of validation*. Paper presented at the International Test Commission conference. Amsterdam, The Netherlands.
13. Lim, G., Buckendahl, C., & Foley, B. (July, 2012). *Evaluating test design through complementary standard setting studies*. Paper presented at the International Test Commission conference. Amsterdam, The Netherlands.
14. Ferdous, A. & Buckendahl, C. (July, 2012). Setting performance standards for partial matrix sampling assessment design. Presentation at the International Testing Commission conference. Amsterdam, The Netherlands.
15. Buckendahl, C. & Burroughs, J. (November, 2011). Business considerations for security: A case study. Presentation at the annual educational conference of the Institute for Credentialing Excellence. New Orleans, LA.
16. Buckendahl, C. & Darragh, I. (November, 2011). Annual reports and renewal best practices. Presentation at the annual educational conference of the Institute for Credentialing Excellence. New Orleans, LA.
17. Buckendahl, C., Koch, L., & Neustel, S. (November, 2011). Pete and repeat: First-time versus repeat candidates and their effect on certification programs. Presentation at the annual educational conference of the Institute for Credentialing Excellence. New Orleans, LA.
18. Buckendahl, C., Davis-Becker, S., Gelin, M., Lorie, W., & Michaels, H. (June, 2011). Can I do this in my pajamas? Validation studies going virtual. Presentation at the annual National Conference on Student Assessment. Orlando, FL.
19. Morgan, D. & Buckendahl, C. (April, 2011). *Evaluating a strategy for setting cut scores on a computer adaptive test*. Paper presented at the annual meeting of the National Council on Measurement in Education. New Orleans, LA.
20. Davis-Becker, S. & Buckendahl, C. (April, 2011). *An empirical evaluation of using domain critical errors in a performance assessment*. Paper presented at the annual meeting of the National Council on Measurement in Education. New Orleans, LA.
21. Ferdous, A. & Buckendahl, C. (April, 2011). *Evaluating panelists' standard setting perceptions in developing nations*. Paper presented at the annual meeting of the National Council on Measurement in Education. New Orleans, LA.
22. Davis-Becker, S. & Buckendahl, C. (April, 2011). *Using the results of multiple standard setting methods to inform policy*. Paper presented at the annual meeting of the National Council on Measurement in Education. New Orleans, LA.

23. Davidson, A., Buckendahl, C. & Betebenner, D. (April, 2011). *Developing a validity framework for establishing expectations for annual student growth*. Paper presented at the annual meeting of the National Conference on Measurement in Education. New Orleans, LA.
24. Buckendahl, C., Davis-Becker, S., & Russell, S. (February, 2011). Defining and developing assessment literacy among stakeholders. Workshop presented at the Association of Test Publishers' Innovations in Testing Conference. Phoenix, AZ.
25. Buckendahl, C., Domaleski, C., Taylor, M., & Willhoft, J. (June, 2010). Improving the coherence of content and performance standards in state assessment systems. Presentation at the annual National Conference on Student Assessment. Detroit, MI.
26. Buckendahl, C., Eberhart, D., Fiandaca, G., & Jensen, J. (June, 2010). Conversations around reporting test results: What are the issues, what's new, what's next? Presentation at the annual National Conference on Student Assessment. Detroit, MI.
27. Buckendahl, C. & Davis, S. (May, 2010). *Evaluating the appropriateness and use of domain critical errors*. Paper presented at the annual meeting of the National Council on Measurement in Education. Denver, CO.
28. Henderson-Montero, D. & Buckendahl, C. (May, 2010). *Psychometric challenges in a rapidly changing environment*. Paper presented at the annual meeting of the National Council on Measurement in Education. Denver, CO.
29. Buckendahl, C. (February, 2010). Using security data to inform legal, policy and psychometric considerations. Presentation at the annual meeting of the Association of Test Publishers. Orlando, FL.
30. Buckendahl, C. (June, 2009). An audit of the NAEP assessment lifecycle. In C. Buckendahl (organizer), Findings from the evaluation of the National Assessment of Educational Progress. Presentation at the annual National Conference on Student Assessment. Los Angeles, CA.
31. Buckendahl, C. (June, 2009). Designing an evaluation of consequences in a state assessment program. In S. Davis (moderator), Developing and evaluating consequential validity evidence for a state assessment program. Presentation at the annual National Conference on Student Assessment. Los Angeles, CA.
32. Buckendahl, C. (April, 2009). *Evaluating rater agreement for credentialing exams with performance items: Integrating policy and practice*. Paper presented at the annual meeting of the National Council on Measurement in Education. San Diego, CA.

33. Davis, S. & Buckendahl, C (April, 2009). *Integration of cognitive demand into credentialing exam development*. Paper presented at the annual meeting of the National Council on Measurement in Education. San Diego, CA.
34. Buckendahl, C. (February, 2009). Reactions to the CBAL project. Presentation at the annual meeting of the Association of Test Publishers. Rancho Mirage, CA.
35. Buckendahl, C. & Davis, S. (July, 2008). *Setting cut scores for a clinical skills licensure test: A comparison of methods*. Paper presented at the International Test Commission Conference. Liverpool, U.K.
36. Buckendahl, C. & Davis, S. (July, 2008). *Communicating the results of a psychometric audit*. Paper presented as part of the symposium, Communicating psychometrics to policymakers: Challenges from the 2004-07 evaluation of the National Assessment of Educational Progress, at the International Test Commission Conference. Liverpool, U.K.
37. Davis, S. & Buckendahl, C. (July, 2008). *The impact of randomly ordering an ordered item booklet in an application of the bookmark standard setting method*. Paper presented at the International Test Commission Conference. Liverpool, U.K.
38. Buckendahl, C. (June, 2008). Considerations when defining the role of contractors in a state assessment program. Presentation at the annual meeting of the Council of Chief State School Officers. Orlando, FL.
39. Davis, S. & Buckendahl, C. (March, 2008). *A proposed framework for evaluating alignment studies*. Paper presented at the annual meeting of the National Council on Measurement in Education. New York, NY.
40. Buckendahl, C., Ferdous, A., & Gerrow, J. (March, 2008). *Setting cut scores with subsets of items: An empirical illustration*. Paper presented at the annual meeting of the National Council on Measurement in Education. New York, NY.
41. Buckendahl, C. (March, 2008). Designing the 2004-08 evaluation of NAEP. Invited presentation to the NAEP Studies Special Interest Group as part of the annual meeting of the American Educational Research Association. New York, NY.
42. Chin, T., Rodeck, E., Buckendahl, C., & Foley, B. (March, 2008). *The impact of consensus on alignment judgments*. Paper presented at the annual meeting of the National Council on Measurement in Education. New York, NY.
43. Badgett, B., Davis, S., & Buckendahl, C. (March, 2008). *Alignment in higher education: Matching assessments to program goals of an online university*. Paper presented at the annual meeting of the National Council on Measurement in Education. New York, NY.

44. Buckendahl, C. (June, 2007). Evaluating computer automated scoring methods for writing assessments. Presentation at the annual meeting of the Council of Chief State School Officers. Nashville, TN.
45. Buckendahl, C. & Davis, S. (April, 2007). Mitigation of item degradation. Presentation at the annual meeting of the Society for Industrial and Organizational Psychologists. New York City, NY.
46. Buckendahl, C. (April, 2007). *Psychometrics in the courtroom: A matter of life and death?* Paper presented at the annual meeting of the National Council on Measurement in Education. Chicago, IL.
47. Davis, S. & Buckendahl, C. (April, 2007). *Evaluating NCLB's peer review process: A comparison of state compliance decisions.* Paper presented at the annual meeting of the National Council on Measurement in Education. Chicago, IL.
48. Irwin, P., Buckendahl, C., & Poggio, A. (April, 2007). *Examinee-centered standard setting: An alternative approach.* Paper presented at the annual meeting of the National Council on Measurement in Education. Chicago, IL.
49. Buckendahl, C. (March, 2007). Evaluating clinical examinations in dentistry and dental hygiene. Presentation at the annual mid-year meeting of the American Association of Dental Examiners. Chicago, IL.
50. Egan, K. & Buckendahl, C. (February, 2007). Philosophies and methods of standard setting. Workshop presented at the annual meeting of the Association of Test Publishers. Palm Springs, CA.
51. Buckendahl, C., Blackhurst, A., & Rodeck, E. (July, 2006). *Adaptation within a language: Considerations for standard setting.* Paper presented at the International Test Commission Conference. Brussels, Belgium.
52. Davis, S., Buckendahl, C., & Plake, B. (July, 2006). *When adaptation is not an option: An application of bilingual standard setting.* Paper presented at the International Test Commission Conference. Brussels, Belgium.
53. Gatti, G. & Buckendahl, C. (April, 2006). *On correctly classifying examinees.* Paper presented at the annual meeting of the American Educational Research Association. San Francisco, CA.
54. Buckendahl, C. (June, 2005). Computer adaptive testing in state assessment (Organizer/Moderator). Presentation at the annual Large Scale Assessment Conference. San Antonio, TX.

55. Buckendahl, C. (April, 2005). *Whose rules? The relationship between the “rules” and “laws” of testing*. Paper presented at the annual meeting of the National Council on Measurement in Education. Montreal, QB.
56. Plake, B., Impara, J., Buckendahl, C., & Ferdous, A. (April, 2005). *Setting multiple performance standards using the yes/no method: An alternative item mapping procedure*. Paper presented at the annual meeting of the National Council on Measurement in Education. Montreal, QB.
57. Buckendahl, C. (July, 2004). NCLB: Psychometric perspectives. Invited presentation at the annual meeting of the American Psychological Association. Honolulu, HI.
58. Plake, B., Buckendahl, C., & Impara, J. (June, 2004). *A classroom-based system for science assessment*. Paper presented at the annual Large Scale Assessment Conference. Boston, MA.
59. Heusman, D. & Buckendahl, C. (June, 2004). Policy implications of requiring sufficiency in local assessment systems. Presentation at the annual Large Scale Assessment Conference. Boston, MA.
60. Buckendahl, C. & Heusman, D. (June, 2004). Sufficiency: Another dimension of alignment. Presentation at the annual Large Scale Assessment Conference. Boston, MA.
61. Buckendahl, C., Huynh, H., Siskind, T., & Saunders, J. (June, 2004). *Vertically moderated standards for South Carolina’s PACT science assessments*. Paper presented at the annual Large Scale Assessment Conference. Boston, MA.
62. Plake, B., Buckendahl, C., & Impara, J. (May, 2004). *Classroom-based assessment system for science: A model*. Paper presented to the National Research Council. Washington, DC.
63. Buckendahl, C., Huynh, H., Siskind, T., & Saunders, J. (April, 2004). *From content standards through technical advisory committee to the State Board of Education: A case study of South Carolina’s 2003 PACT science assessments*. Paper presented as part of an invited symposium at the annual meeting of the National Council on Measurement in Education. San Diego, CA.
64. Buckendahl, C. (February, 2004). Accreditation: Accountability and continuous quality improvement. Presentation at the annual meeting of the Association of Test Publishers. Indian Wells, CA.
65. Buckendahl, C. (May, 2003). Measurement in evidence-based decision making. Invited presentation at the annual meeting of the Canadian Society for Studies in Education. Halifax, NS.

66. Buckendahl, C., Yang, Y., & Ferdous, A. (April, 2003). *An alternative strategy for estimating decision consistency reliability*. Paper presented at the annual meeting of the American Educational Research Association. Chicago, IL.
67. Buckendahl, C., Adams, B., & Goodwin, C. (February, 2003). Experiences with the Buros Accreditation Process. Presentation at the annual meeting of the Association of Test Publishers. Amelia Island, FL.
68. Buckendahl, C., Yang, Y., Juskiewicz, P., & Bhola, D. (June, 2002). *An empirical comparison of computer automated scoring models*. Paper presented at the International Conference on Computer-Based Testing and the Internet. Winchester, England.
69. Buckendahl, C., Christensen, D., & Roschewski, P. (June, 2002). Nebraska's state assessment model: Findings from the first year. Presentation as part of a symposium at the Large Scale Assessment Conference. Palm Desert, CA.
70. Buckendahl, C., Impara, J., & Plake, B. (April, 2002). *A strategy for evaluating district developed assessments for state accountability*. Paper presented at the annual meeting of the American Educational Research Association. New Orleans, LA.
71. Yang, Y., Buckendahl, C., Juskiewicz, P., & Bhola, D. (April, 2002). *An empirical review of strategies for validating computer automated scoring*. Paper presented at the annual meeting of the American Educational Research Association. New Orleans, LA.
72. Boss, T., Endorf, D., & Buckendahl, C. (April, 2002). *Informing state assessment from the local level: A district's reflections*. Paper presented at the annual meeting of the American Educational Research Association. New Orleans, LA.
73. Buckendahl, C., Christensen, D., & Roschewski, P., (June, 2001). Standards, assessment, and accountability – A balanced approach. Presentation as part of a symposium at the Large Scale Assessment Conference. Houston, TX.
74. Buckendahl, C., Impara, J., & Plake, B. (April, 2001). *Computing composite scores for accountability: A validation of Nebraska's district evaluation model*. Paper presented at the annual meeting of the American Educational Research Association. Seattle, WA.
75. Buckendahl, C., Smith, R., Impara, J., & Plake, B. (April, 2001). *A comparison of Angoff and Bookmark standard setting method*. Paper presented at the annual meeting of the National Council on Measurement in Education. Seattle, WA.

76. Plake, B., Buckendahl, C., & Impara, J. (June, 2000). *Alignment of standardized achievement tests to state content standards in reading: A comparison of publishers' and teachers' perspectives*. Paper presented at the Large Scale Assessment Conference. Snowbird, UT.
77. Impara, J., Plake, B., & Buckendahl, C. (June, 2000). *The comparability of nationally norm-referenced tests relative to Nebraska's academic content standards*. Paper presented at the Large Scale Assessment Conference. Snowbird, UT.
78. Buckendahl, C., Impara, J., & Plake, B. (June, 2000). *An alternative strategy for evaluating school district performance for state accountability*. Paper presented at the Large Scale Assessment Conference. Snowbird, UT.
79. Buckendahl, C., Plake, B., Impara, J., & Irwin, P. (April, 2000). *Alignment of standardized achievement tests to state content standards: A comparison of publishers' and teachers' perspectives*. Paper presented at the annual meeting of the National Council on Measurement in Education. New Orleans, LA.
80. Buckendahl, C., Impara, J., Giraud, G., & Irwin, P. (April, 2000). *The consequences of judges making advance estimates of impact*. Paper presented at the annual meeting of the American Educational Research Association. New Orleans, LA.
81. Buckendahl, C., Plake, B., & Impara, J. (April, 1999). *Setting minimum passing scores on high-stakes assessments that combine selected and constructed response formats*. Paper to be presented at the annual meeting of the American Educational Research Association. Montreal, QB.

Regional

1. Buckendahl, C. (January, 2008). Current trends in licensure assessment: Implications for the profession. Presentation at the annual meeting of the Southern Conference of Deans and Dental Examiners. Birmingham, AL.
2. Davis, S. L., Foley, B. P., Buckendahl, C. W., & Chin, T. (October, 2007). *Vertical alignment: Linking content standards across grade levels*. Paper presented at the 2007 annual meeting of the Northern Rocky Mountain Educational Research Association, Jackson Hole, WY.
3. Badgett, B. A., Davis, S. L., & Buckendahl, C. W. (October, 2007). *Alignment in an online university: Matching assessments to program goals*. Paper presented at the 2007 annual meeting of the Northern Rocky Mountain Educational Research Association, Jackson Hole, WY.

4. Badgett, B. & Buckendahl, C. (October, 2006). *Establishing annual measurable achievement objectives for limited English proficiency students: Five options*. Paper presented at the annual meeting of the Northern Rocky Mountain Educational Research Association. Sun Valley, ID.
5. Davis, S. & Buckendahl, C. (October, 2006). *Evaluating the No Child Left Behind peer review process with the Standards for Educational and Psychological Testing*. Paper presented at the annual meeting of the Northern Rocky Mountain Educational Research Association. Sun Valley, ID.
6. Buckendahl, C. (October, 2005). Conducting occupational analyses for licensure testing. Invited presentations to the Public Advocacy Committee and Steering Committee of the North East Regional Board of Dental Examiners, Inc. Washington, DC.
7. Buckendahl, C. (October, 2005). *Challenges to instructionally supportive accountability tests*. Paper presented at the annual meeting of the Northern Rocky Mountain Educational Research Association. Jackson Hole, WY.
8. Buckendahl, C. (October, 2005). Evaluating technical quality for licensure tests. Invited presentation at the annual meeting of the American Association of Dental Examiners. Philadelphia, PA.
9. Buckendahl, C. (October, 2004). Alignment: A rater's perspective. Presentation at the annual meeting of the Northern Rocky Mountain Educational Research Association. Custer, SD.
10. Buckendahl C. & Spies, R. (October, 2004). Developing test evaluation skills: A workshop designed to build assessment literacy. Presentation at the annual meeting of the Northern Rocky Mountain Educational Research Association. Custer, SD.
11. Buckendahl, C. (January, 2004). Evaluating the quality of measurement systems. Invited presentation at the annual meeting of the Southern Conference of Deans and Dental Examiners. Biloxi, MS.
12. Buckendahl, C. (October, 2003). Standards based assessment: Where are we going? Presentation at the annual meeting of the Northern Rocky Mountain Educational Research Association. Jackson Hole, WY.
13. Buckendahl, C. & Yang, Y. (October, 2002). *An alternative strategy for estimating reliability*. Paper presented at the annual meeting of the Northern Rocky Mountain Educational Research Association. Estes Park, CO.
14. Buckendahl, C., Impara, J., & Plake, B. (October, 2001). *A strategy for evaluating district developed assessments for state accountability*. Paper presented at the

- annual meeting of the Mid-Western Educational Research Association. Chicago, IL.
15. Yang, Y., Buckendahl, C., & Juskiewicz, P. (October, 2001). *A review of strategies for validating computer automated scoring*. Paper presented at the annual meeting of the Mid-Western Educational Research Association. Chicago, IL.
 16. Endorf, D., Boss, T., & Buckendahl, C. (October, 2001). *Informing state assessment from the local level: A district's reflections*. Paper presented at the annual meeting of the Mid-Western Educational Research Association. Chicago, IL.
 17. Spies, R., Plake, B., Buckendahl, C., & Impara, J. (October, 2001). *An examination of enhanced teacher feedback in standard setting: Two case studies from the middle school curriculum*. Paper presented at the annual meeting of the Mid-Western Educational Research Association. Chicago, IL.
 18. Buckendahl, C., Plake, B., & Impara, J. (October, 2000). *Computing composite scale scores for accountability: A validation study*. Paper presented at the annual meeting of the Mid-Western Educational Research Association. Chicago, IL.
 19. Buckendahl, C., Smith, R., Impara, J., & Plake, B. (October, 2000). *A comparison of the Angoff and Bookmark standard setting methods*. Paper presented at the annual meeting of the Mid-Western Educational Research Association. Chicago, IL.
 20. Buckendahl, C., Plake, B., Impara, J., Spies, R., Giraud, G., & Irwin, P. (October, 1999). *The feasibility of using multiple assessments to measure student performance on state content standards*. Paper presented at the annual meeting of the Mid-Western Educational Research Association. Chicago, IL.
 21. Colman, R., Wyatt, J., Thompson, R., Wilcox, B., Peterson, R., Jackson, S., Buckendahl, C., Christensen, E. (November, 1998). *Predicting abuse-prone parental attitudes and discipline practices from a national survey*. Paper presented at Midwest Child Maltreatment Conference. Lincoln, NE.

Courses taught

Legal issues in testing (doctoral seminar)

Guest lectures

Intelligence testing: A matter of life and death? (University of Nevada, Las Vegas)
Alignment: Extending content validity theory (University of Massachusetts, Amherst)
Psychometrics in the courtroom (University of Nebraska, Lincoln and University of Massachusetts, Amherst)

A brief history of educational and psychological testing (University of Nebraska, Lincoln)

Legal issues in educational testing (University of Maryland, College Park)

Evaluating technical quality of local assessments (University of Nebraska, Lincoln)

Professional development workshops delivered

Methods for setting cut scores on tests in education and credentialing

Introduction to reliability for local assessments

Introduction to setting mastery levels (standard setting)

Writing multiple choice and constructed response items

Writing items that measure higher order thinking skills

An overview of Nebraska's state assessment system

Understanding validity evidence in local assessment systems

Designing an outcomes assessment system for accreditation

Evaluating alignment of local assessment systems

Consultation

1. National Council of Architectural Registration Boards (NCARB): Provide program design, test design, psychometric consultation, test development and related services. (2011-present).
2. American Board of Emergency Medicine (ABEM): Provide program design, psychometric consultation, test development, data analysis, standard setting, and documentation for a subspecialty certification examination. (2010-present).
3. American Council of Teachers of Foreign Language (ACTFL): Provide standard setting facilitation services for tests of language proficiency. (2011-present).
4. California Bar Examination: Provide ad hoc psychometric consultation and custom data analysis. (2010-present).
5. ALTA Language Services: Provide psychometric consultation, data analysis, and standard setting services for tests of foreign languages. (2008-2011).
6. Comprehensive Adult Student Assessment Systems (CASAS). Provide program design, psychometric consultation, data analysis support, and technical manual development and review. (2008-present).
7. Scrum Alliance: Provide program design and psychometric consultation on characteristics of certification testing program. (2008-2009)
8. Nevada Department of Education: Provide general psychometric and policy consultation on educational assessment programs in reading, mathematics, science, and writing as a member of their Technical Advisory Committee. (2008-present)

9. National Dental Examining Board of Canada (NDEB): Provide psychometric consultation for development, validation, and policy on written and OSCE examination programs for dental licensure. (2006-present)
10. Southern Maine University: Provide program design, psychometric, and legal defensibility consultation for child protective specialist assessments in Maine. (2008-2011)
11. Council of Interstate Testing Agencies (CITA): Provide psychometric consultation on clinical dental and dental hygiene examination development, validation, and policy. (2007-2010)
12. James Satterwhite Academy: Provide psychometric consultation on characteristics of employment testing program and item quality for child protective specialists in New York City. (2006-2008)
13. Federation of State Boards of Physical Therapy (FSBPT): Provide psychometric consultation on exam development, technical quality, and legal defensibility for licensure examinations for physical therapy and physical therapy assistants. (2005-2007)
14. National Council of State Boards of Nursing (NCSBN): Provide psychometric consultation on validity study methodologies including practice analyses and standard setting. (2004-2008)
15. U.S. Department of Education (ED): Principal Investigator for the Congressionally mandated evaluation of the National Assessment of Educational Progress (NAEP). Responsible for management of subcontract with University of Massachusetts, coordination of research activities, coordination of Technical Working Group, and writing/editing reports documenting activities and findings. (2004-2009)
16. American Dental Examinations (ADEX): Provide psychometric consultation on exam development, technical quality, and legal defensibility for proposed national clinical examinations in dentistry and dental hygiene. (2004-2006, 2010-present)
17. Florida Board of Bar Examiners: Provide psychometric consultation on exam development, technical quality, validation, and legal defensibility for a state's bar examination program. (2004-present)
18. North East Regional Board of Dental Examiners (NERB): Provide psychometric consultation on exam development, technical quality, and validation for regional clinical licensure examinations in dentistry and dental hygiene. (2004-2006, 2010-present)

19. Pearson Educational Research: Provide psychometric consultation on alignment studies and computer based classroom assessment methods. (2004-2006)
20. Southern Regional Testing Agency, Inc. (SRTA): Provide psychometric consultation on exam development, technical quality, validation, and legal defensibility for a regional dental and dental hygiene clinical licensure organization. (2002-present)
21. Northwest Evaluation Association (NWEA): Provide psychometric consultation on the technical quality of examinations for adaptive testing programs used in classrooms and state assessment systems. (2002-03)
22. Scantron Corporation: Provide psychometric consultation on the technical quality of examinations for computer adaptive diagnostic classroom assessments. (2002-04)
23. North Carolina State Board of Dental Examiners: Provide psychometric consultation on exam development, technical quality, and validation for one state's dental and dental hygiene licensure examinations. (2002-2004)
24. Certiport, Inc.: Provide psychometric consultation on the technical quality of examinations for information technology certification (IT) program. (2001-03)
25. The Gallup Organization: Provided consultation on appropriate standard setting methods for personnel selection tests. (2001-02)
26. South Dakota Department of Education: Provide psychometric consultation on a state student assessment program. (2001-present)
27. Westside Community Schools: Provide consultation on locally developed assessment program focusing on the technical quality. (2001)
28. Educational Service Units #1, #2, #3, #4, #6, #7, #8, #9, #10, #11, #13, #14, and #19: Conduct workshops and provide consultative services on psychometric quality of assessments for staff developers and member districts across Nebraska. (2000-2006)
29. Galton, a part of Thomson Prometric: Provide psychometric consultation on the technical quality of examinations for information technology (IT) certification programs. (2000-2003)
30. Nebraska Department of Education: Provide measurement and testing consultation services for the state Assessment and Accountability program. (1999-present)

31. Conference for Food Protection: Provide test recognition and consultation services for a national food safety manager certification program. (1999-2002)
32. Oklahoma Commission for Teacher Preparation: Provide psychometric consultation services for a teacher licensure program. (1998-present)
33. Millard Public Schools: Conduct standard setting studies for reading, writing, mathematics, science, and social studies assessments administered in multiple grades. (1998-present)
34. Wichita Public Schools: Provide observation and technical consulting services for standard settings on district-developed tests. (1998-2004)
35. Lincoln Public Schools: Assisted in conducting a standard setting study for high school graduation mathematics examination. (1998-99)
36. National Board for Professional Teaching Standards: Assisted in a content validation study for a national teacher certification program. (1998)
37. YWCA of Lincoln: Conducted program evaluations for the Respite Care Network. (1997-99)

External funding

Contracts/grants awarded as P.I./Co-P.I.: Approximately \$19.9 million since 1999

Contracts/grants not awarded: Approximately \$18.4 million since 1999

Technical and Consulting Reports

Over 190 technical and consulting reports since 1999



The National Center
for the Improvement of
Educational Assessment, Inc.

Vita
CHRIS DOMALESKI
Senior Associate

As a Senior Associate, Chris Domaleski strives to help states and organizations improve student achievement by providing technical support to implement innovative and effective assessment and accountability systems.

Since joining the Center in 2008, Chris has helped multiple states develop, implement, and evaluate assessment and accountability systems. He serves on several technical advisory committees, is the coordinator of the Council of Chief State School Officers State Collaborative on Accountability Systems and Reporting, serves as a technical advisor to two multi-state assessment consortia, and works closely with multiple states to design and validate assessment and accountability systems. Additionally, he has provided technical support to the U.S. Department of Education on multiple occasions to include participation in federal peer review of state of assessment and accountability systems.

Selected current projects include:

- Chair, Technical Advisory Committee, Mississippi Department of Education
- Chair, Technical Advisory Committee, Gwinnett County (Georgia) Public Schools
- Technical Advisory Committee, Kansas Department of Education
- Technical Advisory Committee, South Carolina Department of Education
- Technical Advisory Committee, Nevada Department of Education
- Technical Advisory Committee, Alaska Department of Education
- Technical Advisory Committee, Hawaii Educator Evaluation System
- Consultant to the National Center and State Collaborative (NCSC) Alternative Assessment Consortium for assessment design
- Co-Coordinate/ Facilitate Technical Advisory Committee for the Partnership for the Assessment of Readiness for College and Careers (PARCC) Consortium
- Consultant to the Wyoming Department of Education for Accountability Design
- Consultant to the Utah State Office of Education for Assessment Design and Accountability

Prior to working with the Center, Chris was Associate Superintendent for Assessment and Accountability at the Georgia Department of Education, where he was responsible for the development and administration of the state's K-12 testing program and accountability system. He received his Ph.D. at Georgia State University in Educational Policy Studies, concentrating in Research, Measurement, and Statistics and he currently teaches graduate courses in measurement and statistics at the University of Georgia.

The National Center for the Improvement of Educational Assessment, Inc.
31 Mount Vernon St
Dover, NH 03820
Telephone (603) 516-7900
E-mail cdomaleski@nceia.org
website www.nceia.org

Education

Doctor of Philosophy, Georgia State University, Atlanta, Georgia, 2006
Educational Policy Studies, Concentration in Research, Measurement, and Statistics

Master of Education, The University of Georgia, Athens, Georgia, 1994
Student Personnel in Higher Education

Bachelor of Science, North Georgia College, Dahlonega, Georgia, 1991
Major in Political Science with a minor in English

Dissertation: *Exploring the Efficacy of Pre-Equating a Large-Scale, Criterion-Referenced Assessment with Respect to Measurement Equivalence*

Professional History

National Center for the Improvement of Educational Assessment, Inc.

Dover, New Hampshire, 2008 – Present

Senior Associate

- Work with clients to design and/or validate assessment and accountability systems
- Provide research and consulting services on a variety of policy and/or psychometric issues related to assessment and accountability

Georgia Department of Education, Atlanta, Georgia, 2002 – 2008

Associate Superintendent for Assessment and Accountability, 2007 - 2008

- Responsible for all-aspects of the development and implementation of the state K-12 assessment program and the state accountability system under NCLB
- Oversee the Division of Assessment Research Development, Division of Assessment Administration, and Division of Accountability
- Serve as liaison to the State Superintendent of Schools and the Georgia State Board of Education for policy, programs, and issues related to assessments and accountability

Division Director for Testing, 2006 - 2007

- Responsible for all-aspects of the statewide K-12 assessment program, involving over 4.5 million tests administered annually to over 1 million public school students
- Manage contracts, deliverables, and budgets for multiple state assessment programs
- Ensure state assessment program is in compliance with state and federal laws and regulations

Manager for Research and Development, 2004 – 2006

- Provide leadership for all projects and activities of the Research and Development team for the Testing Division to include: development and validation of statewide assessments, preparation of research reports, management of scoring and reporting, and training and support to assessment stakeholders
- Manage contracts, budgets, and deliverables for multiple projects

Assessment Specialist, 2002 – 2004

- Provide primary psychometric support related to the state assessment programs
- Conduct data analysis to verify the integrity and validity of assessment data.

- Prepare reports for GaDOE and other stakeholders to address assessment properties, research findings, and appropriate use of test scores.

The University of Georgia, Athens, Georgia, 2014- Present

Instructor, College of Education 2014 – Present

- Teach graduate level research and statistics courses
- Courses taught include: Analysis of Variance Methods in Education, Applied Correlation and Regression Methods

Georgia State University, Atlanta, Georgia, 2000 – Present

Instructor, College of Education, 2006 – 2014

- Teach graduate level research, measurement, and statistics courses
- Courses taught include: Research Methods, Institutional Research, Item Response Theory, and Advanced Measurement Theory

Graduate Assistant Educational Research Bureau, 2001 – 2002

- Worked with faculty and graduate students engaged in educational research to provide guidance and support for research design, implementation, statistical analyses, and interpretation
- Supported ongoing research projects of the bureau, including such tasks as developing surveys and assessments, conducting literature reviews, and performing data analysis.

Selected Presentations and Publications

Betebenner, D., Diaz-Bilello, E., Domaleski, C., & Marion, S. (2014). *Student Growth Percentiles During the Assessment Transition: Technical, Practical and Political Implications*. Paper commissioned by the Council of Chief State School Officers.

Bishop, N.S., Sharairi, S., Swift, D., Lei, P., & Domaleski, C.S. (2006, April). *Comparing Growth and AYP Results Over Multiple Years Using Different IRT Scale Transformation Procedures*. Paper presented at the annual meeting of the National Council on Measurement and Education, San Francisco, CA.

Buckendahl, C., Domaleski, C.S., Taylor, M., & Wilhoft, J. (2010, June). *Improving the Coherence of Content and Performance Standards in State Assessment Systems*. Presentation at the annual Council of Chief State School Officers National Conference on Student Assessment, Detroit, MI.

Domaleski, C.S. (2015, June). The Progress and Promise of Open Source Technology to Support State Assessment Systems. Presentation at the annual Council of Chief State School Officers National Conference on Student Assessment, San Diego, CA.

Domaleski, C., Gong, B., Hess, K., Marion, S., Curl, C., Peltzman, A. (2015). *Assessment to Support Competency-Based Pathways*. Washington, DC: Achieve.

Domaleski, C.S. & Hall, E. (2014). *Assessment Transitions and Implications for Accountability*. Paper commissioned by the Council of Chief State School Officers Accountability Systems and Reporting State Collaborative on Assessment and Student Standards.

Domaleski, C.S. & Perie, M. (2012). *Promoting Equity in State Education Accountability Systems*. Dover, NH: NCIEA.

Domaleski, C.S. (2011, April). Considerations for Using Student Performance in Evaluations of Educator Effectiveness. Invited presentation at the Collaborative Conference for Student Achievement. Greensboro, N.C.

Domaleski, C.S. (2011). *State End of Course Tests: A Policy Brief*. Paper commissioned by the Council of Chief State School Officers Technical Issues in Large Scale Assessment State Collaborative on Assessment and Student Standards.

Domaleski, C.S. (2010). Accountability and operational issues. In M. Perie (Ed.), *Teaching and assessing low achieving students with disabilities: A guide to alternate assessments based on modified achievement standards*. Baltimore: Brookes Publishing.

Domaleski, C.S. (2010, June). Comparability: What, Why, When and the Changing Landscape of Computer-Based Testing. Discussant for presentation at the annual Council of Chief State School Officers National Conference on Student Assessment, Detroit, MI.

Domaleski, C.S. (2010, June). Evaluating the Impact of Accountability Systems: State Experiences with Consequential Validity Research. Discussant for presentation at the annual Council of Chief State School Officers National Conference on Student Assessment, Detroit, MI.

Domaleski, C. & Hill, R. (2010). *Considerations for Using Assessment Data to Inform Determinations of Teacher Effectiveness*. Dover, NH: NCIEA.

Domaleski, C.S. (2007, June). Estimating Reliability of Assessment Outcomes for Special Populations: Perspectives, Methods, and Recommendations. Discussant for presentation at the annual Council of Chief State School Officers Large Scale Assessment Conference, Nashville, TN.

Domaleski, C.S., Long, D., Siskind, T., & Reel, M. (2007, June). *Leveraging Technology for State Assessments*. Presentation at the annual Council of Chief State School Officers Large Scale Assessment Conference, Nashville, TN.

Domaleski, C.S. (2007, April). *The Efficacy of Equating with respect to Measurement Equivalence*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Englehard, G., Fincher, M., & Domaleski, C.S. (2011). *Mathematics Performance of Students with and Without Disabilities under Accommodated Conditions Using Resource Guides and Calculators on High Stakes Tests*. *Applied Measurement in Education*, 24(1), 22 – 38.

Erpenbach, W. & Domaleski, C. (2013) *ESEA Flexibility Requests: A Study of States' Requests for Waivers from Requirements of the No Child Left Behind Act of 2001*. Paper commissioned by the Council of Chief State School Officers Accountability Systems and Reporting State Collaborative on Assessment and Student Standards.

Marion, S., DePascale, C., Domaleski, C., Gong, B., & Diaz-Bilello (2012). Considerations for Analyzing Educator's Contributions to Student Learning in Non-Tested Subjects and Grades with a Focus on Student Learning Objectives. Dover, NH: NCIEA.

Neel, J.H., Monaco, M.K., Domaleski, C.S., & Stephens-Bonty, T. (2001, November). *Levene's Test Revisited*. Paper presented at the annual meeting of the Georgia Educational Research Association, Atlanta, GA.

Oshima, T.C., Raju, N.S., & Domaleski, C.S. (2006, April). *Conditional DIF and DTF*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Oshima, T.C. & Domaleski, C.S. (2006, April). Academic Performance Gap Due to Birthdate in Kindergarten Through Eighth Grade. *Journal of Educational Research*.

Patz, R.J., Lall, V., Domaleski, C.S. (2006, April). *Estimating the Rasch Model with Block-Diagonal Item Response Matrix: An Exploration of Winsteps Software with Implications for Equivalent-Groups Equating*. Paper presented at the International Objective Measurement Workshop, Berkeley, CA.

CURRICULUM VITAE -- GEORGE ENGELHARD, JR.

June 27, 2015

PERSONAL INFORMATION

Home Address:

311 Connecticut Ave. NE
Atlanta, GA 30307

email: gengelh@uga.edu

website: www.GeorgeEngelhard.com

University Address:

Educational Psychology, College of Education
The University of Georgia
Athens, GA 30602

EDUCATIONAL BACKGROUND

Ph.D. (1985) in Measurement, Evaluation and Statistical Analysis (MESA program), The University of Chicago, Chicago, Illinois. **Dissertation title:** *The discovery of educational goals and outcomes: A view of the latent curriculum of schooling.* Chair: Professor Benjamin S. Bloom.

M.S. (1977) in Research, Measurement and Evaluation in the Social and Behavioral Sciences, Southern Connecticut State College, New Haven, Connecticut. **Thesis title:** *Dean's list recognition and subsequent academic performance.*

B.A. (1975) in Sociology with a minor in History, Southern Connecticut State College, New Haven, Connecticut.

PROFESSIONAL EXPERIENCE

Professor, The University of Georgia, College of Education. September 2013 to present.

Program Coordinator, Quantitative Methodology Program, Educational Psychology, The University of Georgia, College of Education. September 2013 to present.

Professor Emeritus, Emory University, Division of Educational Studies. September 2013 to present.

Professor, Emory University, Division of Educational Studies. September 1997 to August 2013.

Director of Graduate Studies, Emory University, Division of Educational Studies, (Fall 2011 to Spring 2013)).

Associate Professor, Emory University, September 1991 to August 1997.

Assistant Professor, Emory University, September 1985 to August 1991.

Assistant Director of the Office of Institutional Research and Evaluation at Chicago State University. September 1980 to August 1985.

BOOKS

Engelhard, G. (in preparation). *Invariant measurement with raters: Classical and modern measurement theories for rating scales*. New York: Routledge.

Engelhard, G. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences*. New York: Routledge.

EDITED VOLUMES

Garner, M., Engelhard, G., Wilson, M., & Fisher, W. (Eds.) (2010). *Advances in Rasch Measurement, Volume One*. Maple Grove, MN: JAM Press

Wilson, M., & Engelhard, G. (Eds.). (2000). *Objective Measurement: Theory into Practice, Volume 5*. Stamford, CT: Ablex.

Wilson, M., Engelhard, G., & Draney, K. (Eds.). (1997). *Objective Measurement: Theory into Practice, Volume 4*. Norwood, NJ: Ablex.

Engelhard, G., & Wilson, M. (Eds.). (1996). *Objective Measurement: Theory into Practice, Volume 3*. Norwood, NJ: Ablex.

FELLOWSHIPS, GRANTS, AND AWARDS

Fellow of the American Education Research Association (2009 to present).

Visiting Research Scientist, The College Board, New York City (Fall 2012).
Sabbatical leave at the College Board to continue my research and policy collaborations with colleagues at the College Board.

Visiting Scholar at Center for English Language Communication, The National University of Singapore (January 2012). The purpose of the visiting scholars program is to create opportunities for collaborative work between the CELC Centre and internationally recognized and distinguished scholars.

Visiting Scholar, The College Board, New York City (Fall 2010). The purpose of this visiting scholarship is to provide opportunities for meaningful research and policy interactions among distinguished research scholars in academia and members of the College Board.

National Academy of Education Travel Scholarship (2010 Fellows Forum) to attend National Academy of Education Annual Meeting (Washington, DC, November 2010).

Distinguished Lecturer at Tokyo and Osaka Centers (December 2007, December 2009). The purpose of this lecture series is to bring to Japan internationally recognized scholars for intensive weekend seminars on issues related to English as a second language and applied linguistics.

Halle Faculty Delegation to Germany (June 4-17, 2005). The purpose of this program was to provide Emory faculty with a cultural and scholarly introduction to contemporary

political, business, academic, and artistic issues in Germany and the European Union.

Atlantik-Brucke Fellow (June 2005 to present). The Atlantik-Brucke (Atlantic Bridge) was founded in 1952 as a private, non-partisan organization that seeks to strengthen ties between Germany and the United States.

Emory Faculty Language and Culture Program in Salamanca, Spain (June 1- July 3, 2001). This program provided an intensive 4-week Spanish Language Course, Seminar on Hispanic Cultural Literacy, and the opportunity for collaborative scholarship at the Universidad de Salamanca.

The College Board, 1999 to 2001. Project title: Monitoring reader performance in the Advanced Placement Program with a many-faceted Rasch model.

The College Board, 2001-2004. Project title: Monitoring reader performance and DRIFT in the AP English Literature and Composition Exam using benchmark essays.

The National Board for Professional Teaching Standards (NBPTS), 1999 to 2001. Project title: Investigating assessor effects in NBPTS assessments for Early Childhood/Generalist and Middle Childhood/Generalist Certification.

Spencer Fellow, National Academy of Education, 1987-89. Project title: The three disciplines of measurement in the behavioral sciences.

Lilly Fellow, Lilly Post-Doctoral Teaching Award, Emory University, 1986-87. Project: Development of a new undergraduate course for Emory College on *Education as a Social Science*. Syllabus selected for inclusion in the American Sociological Association's publication on teaching sociology of education.

Sears Fellow, Writing Across the Curriculum Project, Emory University, 1989-1990. Project: Enhancement of the writing skills of college students.

Distinguished Research Award (March 1992), (with Gordon, B., Siddle Walker, E.V., & Gabrielson, S.). Influences of writing tasks and gender on the quality of writing for black and white students. Eastern Educational Research Association.

Distinguished Paper Award (October 1990), (with J.A. Monsaas). Attitudes toward testing practices as cheating as predictors of teachers' test preparation and test administration practices. Georgia Educational Research Association.

Distinguished Paper Award (November 1989), (with C. Tocci). Mathematics achievement, parental support and gender differences in attitudes toward mathematics of American and Thai adolescents. Georgia Educational Research Association.

Spencer Foundation, Small Grants Program (1994-95). Project title: The elicitation and analysis of expert judgment in educational assessment.

Summer Grant, Emory Center for Teaching and Curriculum (1996). Summer stipend to

support the development of new course for Emory College entitled: An introduction to educational research.

Member, (February 1992 to August 1996). Measurement Research Team (MRT) for National Board of Professional Teaching Standards, Center for Educational Research and Evaluation at the University of North Carolina at Greensboro. MRT was responsible for defining a measurement research agenda that is consistent with the NBPTS goal of developing a technically sound system of assessing and certifying highly accomplished classroom teachers.

Regular Grant, University Research Committee, (September 1990 to August 1991), Emory University. Project title: Edward L. Thorndike, Ben D. Wood, and the emergence of two research traditions in educational measurement, 1904-1923.

Summer Grant, University Research Committee, (April 1991 to August 1991), Emory University. Project title: Major research traditions in the measurement of individual differences: An emergent history.

Summer Research Fellowship, (June 1989). *Secondary Analyses of the Second International Mathematics Study Data*, National Science Foundation. University of Illinois at Urbana-Champaign.

Faculty Development Award, Emory University. Project: Using Computer Assisted Instruction in College Classrooms (Spring 1992).

Summer Faculty Development Awards, Emory University, 1986 and 1990.

Chicago State University Foundation. Research grant (April 1982). Project title: The Latent Curriculum of Schooling.

Tuition Scholarships, The University of Chicago, Fall, Winter and Spring Quarters (1978, 1979, 1980).

PUBLICATIONS

Wesolowski, B., Wind, S.A., & Engelhard, G. (2015). Rater Fairness in Music Performance Assessment: Evaluating Model-Data Fit and Differential Rater Functioning. *Musicae Scientiae*, 147-170.

Walker, A.A., & Engelhard, G. (in press). Exploring Person Fit with an Approach Based on Multilevel Logistic Regression. *Applied Measurement in Education*.

Chang, M.L., & Engelhard, G. (in press). Examining Teacher Sense of Efficacy Scale at the Item Level With Rasch Measurement Model. *Journal of Psychoeducational Assessment*.

Ricketts, S., Chang, M.L., & Engelhard, G. (2015). Development and Validation of a Scale to Measure Academic Resilience in Mathematics. *European Journal of Psychological Assessment*

Behizadeh, N. & Engelhard, G. (2015). Involving Diverse Communities of Practice to Minimize Unintended Consequences of Test-Based Accountability Systems. *Measurement: Interdisciplinary Research & Perspectives*, 13, 26-30.

Behizadeh, N. & Engelhard, G. (in press). What is a valid writing assessment from the perspectives of the writing and measurement communities? *Pensamiento Educativo: Revista de Investigación Educativa Latinoamericana*.

Razynski, K., Engelhard, G., Cohen, A., & Lu, Z. (in press). Comparing the Effectiveness of Self-Paced and Collaborative Frame-of-Reference Training on Rater Accuracy in a Large Scale Writing Assessment. *Journal of Educational Measurement*.

Walker, A.A., Engelhard, G., Royal, K., & Hedgpeth, M. (in review). Exploring Aberrant Responses Using Person Fit and Person Response Functions. *Journal of Applied Measurement*.

Wesolowski, B., Wind, S.A., & Engelhard, G. (under revision). Examining Rater Behaviors and Rating Scale Structure in Music Performance Assessment. *Music Perception*.

Wind, S. A. & Engelhard, G., Jr. (in review). Evaluating rating quality in writing assessment using Mokken scale analysis. *Educational and Psychological Measurement*.

Engelhard, G., & Wang, J. (2015). Unfolding rater accuracy in performance assessments. *Rasch Measurement Transactions*, 28(4), 1489-1491.

Wang, J., Engelhard, G., & Lu, Z. (2014). Clarifying the Conceptualization of Indicators Within Different Models. *Measurement: Interdisciplinary Research & Perspectives*, 12, 155-159.

Engelhard, G., Jr., Kobrin, J. L., & Wind, S. A. (2014). Exploring differential subgroup functioning on SAT writing items: What happens when English is NOT a test-taker's best language? *International Journal of Testing*, 14: 339–359.

Wind, S. A., & Engelhard, G., Jr. (2014). Model-data fit and adjustments for rater effects. *Rasch Measurement Transactions*, 28(3), 1483-4.

Engelhard, G., & Wang, J. (2014). Alternative Measurement Paradigms for Measuring Executive Functions: SEM (Formative and Reflective Models) and IRT (Rasch Models). *Measurement: Interdisciplinary Research & Perspectives*, 12: 102–108,

Wang, J., & Engelhard, G. (2014). A Pairwise Algorithm in R for Rater-Mediated Assessments. *Rasch Measurement Transactions*, 28(1), 1457-1459.

Walker, A.A., & Engelhard, G. (2014). Game-Based Assessments: A Promising Way to Create Idiographic Perspectives. *Measurement: Interdisciplinary Research & Perspectives*. 12: 57–61

Behizadeh, N. & Engelhard, G. (2014). Development and Validation of a Scale to Measure Perceived Authenticity in Writing. *Assessing Writing*, 21, 18-36.

Walker, E.R., Engelhard, G., Barmon, C., McGee, R.E., Sterk, C.E., DiIorio C., Thompson, N.J. (2014). A mixed methods analysis of support for self-management behaviors: Perspectives of people with epilepsy and their support providers. *Epilepsy & Behavior, 31*, 152–159.

Engelhard, G., & Perkins, A. (2014). Review of BEST Literacy Test. In J.F. Carlson, K.F. Geisinger, & J.L. Jonson (Eds.). *Nineteenth Mental Measurements Yearbook* (pp. 84-85). Lincoln, NE: The University of Nebraska Press.

Engelhard, G., & Perkins, A. (2014). Review of Green's Emotional Perception Test. In J.F. Carlson, K.F. Geisinger, & J.L. Jonson (Eds.). *Nineteenth Mental Measurements Yearbook* (pp. 326-327). Lincoln, NE: The University of Nebraska Press.

Wind, S.A., & Engelhard, G. (2013). How Invariant and Accurate are Domain Ratings in Writing Assessment? *Assessing Writing, 18*, 278–299.

Engelhard, G., & Perkins, A. (2013). Goodness of Model-Data Fit and Invariant Measurement. *Measurement: Interdisciplinary Research & Perspectives, 11*, 112-116.

Engelhard, G. (2013). Hanning person response functions. *Rasch Measurement Transactions, 26*(4), 1392-1393.

Engelhard, G., & Wind, S.A. (2013). Educational testing and schooling: Unanticipated consequences of purposive social action. *Measurement: Interdisciplinary Research and Perspectives, 11*, 1-6.

Kaliski, P., Engelhard, G., Morgan, D., Reshetar, R., Wind, S.A., & Plake, B. (2013) Using the Many-Facet Rasch Model to Evaluate Standard-Setting Judgments: Setting performance standards for Advanced Placement examinations. *Educational and Psychological Measurement, 73*, 386-411

Reisinger, E., Engelhard, G., & Thompson, N.J. (2012). Using Rasch measurement theory to assess three depression scales among adults with epilepsy. *Seizure: European Journal of Epilepsy, 21*, 437-443.

Engelhard, G., Fincher, M., & Domaleski, C. S. (2010). Mathematics performance of students with and without disabilities under accommodated conditions using resource guides and calculators on high stakes tests. *Applied Measurement in Education, 24*(1), 22–38.

Wind, S.A., & Engelhard, G. (2012). Evaluating the quality of ratings in writing assessment: Rater agreement, precision, and accuracy. *Journal of Applied Measurement, 13*(4), 321-335.

Caines, J., & Engelhard Jr., G. (2012). How good is good enough?: Educational standard setting and its effects on African American test takers. *Journal of Negro Education 81*(3), 228-240.

Engelhard, G. (2012). Rasch measurement theory and factor analysis. *Rasch*

Measurement Transactions, 26(3), 1375.

Wind, S.A., & Engelhard, G. (2012). Mapping differential item functioning (DIF maps). *Rasch Measurement Transactions*, 26(1), 1356-1357.

Engelhard, G., & Behizadeh, N. (2012). Exploring the alignment of writing self-efficacy with writing achievement using Rasch measurement theory and qualitative methods. *Journal of Applied Measurement*, 13(2), 1-14.

Engelhard, G., & Behizadeh, N. (2012). Epistemic iterations and consensus definitions of validity. *Measurement: Interdisciplinary Research and Perspectives*, 10(1), 55-58.

Engelhard, G. (2011). Evaluating the Bookmark Judgments of Standard-Setting Panelists. *Educational and Psychological Measurement*, 71(6), 909-924.

Randall, J. & Engelhard, G. (2011). Using Guttman's Facet Theory and Rasch Measurement Theory to examine teacher grades. In N.J.S. Brown, B. Duckor, K. Draney, and M. Wilson (Eds.) *Advances in Rasch Measurement, Volume Two* (pp. 296-326). Chicago: JAM Press.

Engelhard, G., & Perkins, A.F. (2011). Person response functions and the definition of units in the social sciences. *Measurement: Interdisciplinary Research and Perspectives*, 9, 40-45.

Hamm, C., Schulz, M., & Engelhard, G. (2011). Standard setting for the National Assessment of Educational Progress: Evidence regarding the transition from Angoff-based to Bookmark-based methods. *Educational Measurement: Issues and Practice*, 30(2), 3-14.

Perkins, A., Quaynor, L., & Engelhard, G. (2011). The Influences of Home Language, Gender, and Social Class on Mathematics Literacy in France, Germany, Hong Kong, and the United States. *IEA-ETS Research Institute (IERI), Monograph Series*, 35-38.

Randall, J., Cheong, Y.F., & Engelhard, G. (2011). Using Explanatory IRT Modeling to Investigate Context Effects of Differential Item Functioning for Students with Disabilities. *Educational and Psychological Measurement*, 71(1), 129-147.

Behizadeh, N. & Engelhard, G. (2011). Historical view of the influences of measurement and writing theories on the practice of writing assessment. *Assessing Writing*, 16, 189-211.

Engelhard, G., & Sullivan, R.K. (2011). An Ecological Perspective on Learning Progressions as Road Maps for Learning. *Measurement: Interdisciplinary Research and Perspectives*, 9: 138-145, 2011.

Vogel, S.P., & Engelhard, G. (2011) Using Rasch Measurement Theory to Examine Two Instructional Approaches for Teaching and Learning of French Grammar. *Journal of Educational Research*, 104, 267-282.

Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers.

Teaching and Teacher Education, 26(7), 1372-1380.

Randall, J., Engelhard, G., Fincher, M., & Domaleski, D.S. (2010). Performance of Students With and Without Disabilities Under Modified Conditions: Using Resource Guides and Read Aloud Test Modifications on a High Stakes Reading Test. *Journal of Special Education*, 44(2), 79-93.

Engelhard, G. (2010). Review of *Test of Understanding in College Economics, Fourth Edition*. In K.F. Geisinger & R.A. Spies (Eds.). *Eighteenth Mental Measurements Yearbook*. (pp.654-656) . Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (2010). Review of *Diagnostic Test for Pre-Algebra Mathematics*. In K.F. Geisinger & R.A. Spies (Eds.). *Eighteenth Mental Measurements Yearbook* (pp.191-193). Lincoln, NE: The University of Nebraska Press.

Randall, J., & Engelhard, G. (2010). Differences in the grading practices of elementary and middle school teachers. *Journal of Educational Research*, 102(3), 175-186.

Randall, J. & Engelhard, G. (2010). Examining the grading practices of teachers, *Teaching and Teacher Education*, 26, 1372-1380.

Randall, J., & Engelhard, G. (2010). Using Confirmatory Factor Analysis and Rasch Measurement Theory to Assess Measurement Invariance in a High Stakes Reading Assessment. *Applied Measurement in Education*, 23, 286-306.

Garner, M., & Engelhard, G. (2010). Extension of the pairwise algorithm to the rating scale and partial credit models. In Garner, M., Engelhard, G., Wilson, & M., Fisher, W. (Eds.). *Advances in Rasch Measurement, Volume One* (pp. 45-63). JAM Press

Gyagenda, I.S., & Engelhard, G. (2010). Rater, domain, and gender influences on the assessed quality of student writing. In Garner, M., Engelhard, G., Wilson, & M., Fisher, W. (Eds.). *Advances in Rasch Measurement, Volume One* (pp. 398-429). JAM Press

Engelhard, G., & Myford, C. (2010). Comparison of single and double assessor scoring designs for the assessment of accomplished teaching. In Garner, M., Engelhard, G., Wilson, M., & Fisher, W. (Eds.). *Advances in Rasch Measurement, Volume One* (pp. 342-368). Maple Grove, MN: JAM Press.

Randall, J., & Engelhard, G. (2010). Using Guttman's mapping sentences and many-facet Rasch measurement theory to develop an instrument that examines the grading philosophies of teachers. *Journal of Applied Measurement*, 11(2), 1-20.

Engelhard, G. (2009). Evaluating the judgments of standard-setting panelists using Rasch Measurement Theory. In E. V. Smith, Jr., and G. E. Stone (Eds.), *Criterion referenced testing: Practice analysis to score reporting using Rasch measurement models* (pp. 312-346). Maple Grove, MN: JAM Press.

Engelhard, G. (2009). Using IRT and Model-Data fit to Conceptualize Differential Item

and Person Functioning for Students with Disabilities. *Educational and Psychological Measurement*, 69(4), 585-602.

Perkins, A., & Engelhard, G. (2009). Crossing person response functions. *Rasch Measurement Transaction*, 23(1), 1183-1184.

Gyagenda, I.S., & Engelhard, G. (2009). Using classical and modern measurement theories to explore rater, domain, and gender Influences on student writing ability. *Journal of Applied Measurement*, 10(3), 225-246.

Garner, M., & Engelhard, G. (2009). Using paired comparison matrices to estimate parameters of the partial credit Rasch measurement model for rater-mediated assessments. *Journal of Applied Measurement*, 10(1), 30-41.

Engelhard, G., & Myford, C. (2009). Comparison of single and double assessor scoring designs for the assessment of accomplished teaching. *Journal of Applied Measurement*, 10(1), 52-69.

Randall, J., & Engelhard, G. (2009). Examining teacher grades using Rasch measurement theory. *Journal of Educational Measurement*, 46(1) 1-18.

Engelhard, G. (2008). Tuneable Goodness-of-Fit Statistics. *Rasch Measurement Transactions* 22(1), 1158-1159.

Engelhard, G. (2008). Evaluating Body-of-Work Judgments of Standard-Setting Panelists. *Rasch Measurement Transactions*, 21(4).

Engelhard, G. (2008). Historical perspectives on invariant measurement: Guttman, Rasch, and Mokken [Focus article]. *Measurement: Interdisciplinary Research and Perspectives* (6), 1-35.

Engelhard, G. (2008). Rejoinder: Continuing the dialogue on invariant measurement. *Measurement: Interdisciplinary Research and Perspectives*, (6), 1-6.

Engelhard, G. (2008). Standard Errors for Performance Standards based on Bookmark Judgments. *Rasch Measurement Transaction*, 22(1), 1156-1157.

Engelhard, G. (2008). Differential rater functioning. *Rasch measurement: Transactions of the Rasch measurement SIG* (p. 1124), American Educational Research Association.

Alviar-Martin, T., Randall, J., Usher, E.L., & Engelhard, G. (2008). Teaching civic topics in four societies: Examining national context and teacher confidence. *The Journal of Educational Research*. 101, 177-188.

Wolfe, E.W., Myford, C.M., Engelhard, G., & Manalo, J.R. (2007). *Monitoring reader performance and DRIFT in the AP English Literature and Composition Examination using benchmark essays*. New York: The College Board.

Engelhard, G. (2007) Review of *Iowa Test of Basic Skills, Forms A and B*. In K.F.

Geisinger, R.A. Spies, J.F. Carlson, & B.S. Plake (Eds.). *Seventeenth Mental Measurements Yearbook*. . Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (2007) Review of *Reading Fluency Indicator*. In K.F. Geisinger, R.A. Spies, J.F. Carlson, & B.S. Plake (Eds.). *Seventeenth Mental Measurements Yearbook*. . Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (2007). Evaluating bookmark judgments. *Rasch measurement: Transactions of the Rasch measurement SIG*, American Educational Research Association, 21(2), 1097-1098.

Engelhard, G., and Sullivan, R. (2007). Re-conceptualizing validity within the context of a new measure of mathematical knowledge for teaching. *Measurement: Interdisciplinary Research and Perspectives*, 5(2-3), 142-155.

Engelhard, G. (2006). Using time wisely. In T. Guskey (Ed.). *Benjamin S. Bloom: Portraits of an Educator* (pp.16-17) Oxford: Rowman & Littlefield Education.

Engelhard, G. (2006). Not every effort is a success. In T. Guskey (Ed.). *Benjamin S. Bloom: Portraits of an Educator* (pp.60-61) Oxford: Rowman & Littlefield Education.

Engelhard, G. (2005) Review of *Comprehensive Mathematical Abilities Test*. In J. Impara & B. Plake (Eds.), *Sixteenth Mental Measurements Yearbook*. . Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (2005) Review of *Wide Range Achievement Test-Expanded Edition*. In J. Impara & B. Plake (Eds.), *Sixteenth Mental Measurements Yearbook*. Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (2005). Book Review of *Analyzing rater agreement: Manifest variable methods* by Alexander von eye and Eun Young Mun. *Applied Psychological Measurement*.

Engelhard, G. (2005). Guttman Scaling. In K. Kempf-Leonard (Ed.), *Encyclopedia of Social Measurement, Volume 2* (pp. 167-174). San Diego, CA: Academic Press (Elsevier Science).

Engelhard, G. (2005). Item Response Theory (IRT) models for rating scale data. *Encyclopedia of Statistics in Behavioral Science (Vol. 2)*, pp. 995-1003.

Bazarova, S.I., & Engelhard, G. (2004). Comparisons of the mathematics achievement of eighth grade students in the United States and the Russian Federation. *Journal of Applied Measurement*, 5(2), 142-159.

Engelhard, G. (2003). Commentary on T.M. Porter's *Trust in Numbers: The pursuit of objectivity in science and public life*. *Measurement: Interdisciplinary Research and Perspectives*. Website: <http://bear.soe.berkeley.edu/measurement/pubs/toc14.html>.

Engelhard, G. (2003) Review of *The MaroonDAH Assessment Profile for Problem*

Gambling. In J. Impara & B. Plake (Eds.), *Fifteenth Mental Measurements Yearbook*. . Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (2003) Review of *Search Institute Profiles of Student Life: Attitudes and Behaviors*. In J. Impara & B. Plake (Eds.), *Fifteenth Mental Measurements Yearbook*. Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (2003). Review of *Applying the Rasch model: Fundamental measurement in the human sciences* by T.G. Bond and C.M. Fox. *Journal of Educational Measurement*, 40(2), 185-187.

Engelhard, G. & Myford, C.M. (2003). *Monitoring faculty consultant performance in the Advanced Placement English Literature and Composition Program with a many-faceted Rasch model*. New York: College Entrance Examination Board.

Engelhard, G. (2002). Monitoring raters in performance assessments. In G. Tindal and T. Haladyna (Eds.), *Large-scale Assessment Programs for ALL Students: Development, Implementation, and Analysis*, (pp. 261-287). Mahwah, NJ: Erlbaum.

Myford, C.M., & Engelhard, G. (2002). Evaluating the psychometric quality of the National Board for Professional Teaching Standards Early Childhood/Generalist Assessment System, *Journal of Personnel Evaluation in Education*, 15(4), 253-285.

Garner, M., & Engelhard, G. (2002). An eigenvector method for estimating item parameters of the dichotomous and polytomous Rasch models. *Journal of Applied Measurement*, 3(2), 107-128

Hambleton, R.K., Plake, B.S., & Engelhard, G. (2001, summer). Richard M. Jaeger's contributions to the field of standard setting. *Newsletter of Division D (Measurement and Research Methodology)*, *American Educational Research Association*, 11(1).

Engelhard, G. (2001) Review of *Closed High School Placement Test*. In J. Impara & B. Plake (Eds.), *Fourteenth Mental Measurements Yearbook*. (pp. 272-274). Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (2001) Review of *Social Phobia and Anxiety Inventory*. In J. Impara & B. Plake (Eds.), *Fourteenth Mental Measurements Yearbook* (pp. 1161-1163). Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (2001). Historical views of the influences of measurement and reading theories on the assessment of reading. *Journal of Applied Measurement*, 2(1), 1-28.

Engelhard, G., Myford, C., & Cline, F. (2000). *Investigating assessor effects in National Board for Professional Teaching Standards Assessments for Early Childhood/Generalist and Middle Childhood/Generalist Certification*. Princeton, NJ: Educational Testing Service, Research Report No: RR-00-13.

Engelhard, G., & Gordon, B. (2000). Setting and evaluating performance standards for

high stakes writing assessments. In M. Wilson & G. Engelhard (Eds.), *Objective Measurement: Theory into Practice, Volume 5* (pp. 3-14). Stamford, CT: Ablex.

Garner, M., & Engelhard, G. (2000). Rasch measurement theory: The method of paired comparisons and graph theory. In M. Wilson & G. Engelhard (Eds.), *Objective Measurement: Theory into Practice, Volume 5* (pp. 259-286). Stamford, CT: Ablex.

Engelhard, G., Davis, M., & Hansche, L. (1999). Evaluating the accuracy of judgments obtained from item review committees. *Applied Measurement in Education, 12*(2), 199-210.

Garner, M., & Engelhard, G. (1999). Gender differences in performance on multiple-choice and constructed response mathematics items. *Applied Measurement in Education, 12*(1), 29-51.

Engelhard, G. (1998) Review of *CTB Writing Assessment System*. In J. Impara & B. Plake (Eds.), *Thirteenth Mental Measurements Yearbook* (pp. 329-331). Lincoln, NE: The University of Nebraska Press.

Engelhard, G. (1998) Review of *Stanford Diagnostic Reading Test, Fourth Edition* In J. Impara & B. Plake (Eds.), *Thirteenth Mental Measurements Yearbook* . (pp. 939-941). Lincoln, NE: The University of Nebraska Press.

Engelhard, G., & Stone, G.E. (1998). Evaluating the quality of ratings obtained from standard-setting judges. *Educational and Psychological Measurement, 58*(2), 179-196.

Engelhard, G., & Anderson, D.W. (1998). A binomial trials model for examining the ratings of standard-setting judges. *Applied Measurement in Education, 11*(3), 209-230.

Engelhard, G., & Cramer, S. (1997). Using Rasch Measurement to evaluate the ratings of standard-setting judges. In M. Wilson, G. Engelhard, & K. Draney. (Eds.). *Objective Measurement: Theory into Practice, Volume 4* (pp. 97-112). Norwood, NJ: Ablex.

Engelhard, G. (1997). Constructing rater and task banks for performance assessments. *Journal of Outcome Measurement, 1*(1), 19-33.

Engelhard, G. (1997). (Ed.). Introduction to special issue on history of measurement theory. *Educational Measurement: Issues and Practice, Summer, 5-7*.

Engelhard, G. (1996). Evaluating rater accuracy in performance assessments. *Journal of Educational Measurement, 33*(1), 56-70.

Engelhard, G. (1996). Clarification to "Examining rater errors in the assessment of written composition with a many-faceted Rasch model." *Journal of Educational Measurement, 33* (1), 115-116.

Loyd, B., Engelhard, G., & Crocker, L. (1996). Achieving form to form comparability: Fundamental issues and proposed strategies for equating performance assessments of teachers. *Educational Assessment, 3* (1), 99-110.

Monsaas, J.A., & Engelhard, G. (1996). Examining changes in the home environment with the Rasch measurement model. In G. Engelhard, & M. Wilson (Eds.), *Objective Measurement: Theory into Practice, Volume 3* (pp. 127-140). Norwood, NJ: Ablex.

Gordon, B., Engelhard, G., Gabrielson, S., & Bernknopf, S. (1996). Conceptual issues in equating performance assessments: Lessons from writing assessment. *Journal of Research and Development in Education*, 29 (2), 81-88.

Engelhard, G. (1995) Review of *PERSONALYSIS*. In J.C. Conoley & J. Impara (Eds.), *Twelfth Mental Measurements Yearbook* (pp. 768-769). Lincoln, NE. The University of Nebraska Press.

Engelhard, G. (1995). Review of *Entrance Examination for Schools of Practical/Vocational Nursing*. In J.C. Conoley & J. Impara (Eds.), *Twelfth Mental Measurements Yearbook* (pp. 361-362). Lincoln, NE. The University of Nebraska Press.

Gabrielson, S.J., Gordon, B., & Engelhard, G. (1995). The effects of task choice on the quality of writing obtained in a statewide assessment. *Applied Measurement in Education*, 8(4), 273-290.

Engelhard, G. (1994). Examining rater errors in the assessment of written composition with a many-faceted Rasch model. *Journal of Educational Measurement*, 31(2), 93-112.

Linacre, J.M., Engelhard, G., Tatum, D.S., & Myford, C.M. (1994). Measurement with judges: Many-faceted conjoint measurement. *International Journal of Educational Research*, 21(6), 569-577.

Engelhard, G. (1994). Historical views of the concept of invariance in measurement theory. In M. Wilson (Ed.), *Objective Measurement: Theory into Practice, Volume 2* (pp. 73-99). Norwood, NJ: Ablex.

Engelhard, G., Gordon, B., Siddle Walker, E.V., & Gabrielson, S. (1994). Writing tasks and gender: Influences on writing quality of black and white students. *Journal of Educational Research*, 87(4), 197-209.

Monsaas, J.A., & Engelhard, G. (1994). Teachers' attitudes toward testing practices. *Journal of Psychology*, 128(4), 469-477.

Weinburgh, M.H., & Engelhard, G. (1994). Gender, prior academic performance and beliefs as predictors of attitudes toward biology laboratory experiences. *School Science and Mathematics*, 94(3), 118-123.

Engelhard, G. (1994). Resolving the attenuation paradox. *Rasch Measurement Transactions*, 8(3), 379.

Engelhard, G. (1993). Reactions to the attenuation paradox. *Rasch Measurement Transactions*, 7(2), 294.

Engelhard, G. (1993). What is the attenuation paradox? *Rasch Measurement Transactions*, 6(4), 257.

Engelhard, G. (1992). The measurement of writing ability with a many-faceted Rasch model. *Applied Measurement in Education*, 5(3), 171-191.

Engelhard, G., Gordon, B., & Gabrielson, S. (1992). The influences of mode of discourse, experiential demand, and gender on the quality of student writing. *Research in the Teaching of English*, 26(3), 315-336.

Monsaas, J.A., & Engelhard, G. (1992). Home environment and the competitiveness of highly accomplished individuals in four talent fields. In R.S. Albert (Ed.), *Genius and Eminence*. Oxford: Pergamon Press. [Reprint of 1990 article in *Developmental Psychology*.]

Johnson, C., & Engelhard, G. (1992). Gender, academic achievement, and preferences for cooperative, competitive, and individualistic learning among African-American adolescents. *Journal of Psychology*, 126(4), 385-392.

Engelhard, G. (1992). Historical views of invariance: Evidence from the measurement theories of Thorndike, Thurstone and Rasch. *Educational and Psychological Measurement*, 52(2) 275-292.

Engelhard, G. (1992). Anomaly, paradox and progress. *Rasch Measurement Transactions*, 6(2), 212.

Engelhard, G. (1992). Progress in measurement theory. *Rasch Measurement Transactions*, 6(1), 204.

Engelhard, G. (1992). Thorndike's scaling vs. Wood's scoring. *Rasch Measurement Transactions*, 5(4), 182.

Engelhard, G. (1992). Review of *California Psychological Inventory: Revised Edition*. In J.J. Kramer & J.C. Conoley, (Eds.), *Eleventh Mental Measurements Yearbook*, (pp. 139-141). Lincoln, NE: University of Nebraska Press.

Engelhard, G. (1992). Review of *The Psychap Inventory*. In J.J. Kramer & J.C. Conoley, (Eds.), *Eleventh Mental Measurements Yearbook*, (pp. 725-727). Lincoln, NE: University of Nebraska Press.

Engelhard, G. (1991). Thorndike, Thurstone and Rasch: A comparison of their approaches to item-invariant measurement. *Journal of Research and Development in Education*, 24(2), 45-60.

Tocci, C.M., & Engelhard, G. (1991). Achievement, parental support and gender differences in attitudes toward mathematics. *Journal of Educational Research*, 84, 280-286.

Engelhard, G. (1991). The peer review process in Division D. *d'News*, Newsletter of Division D (Measurement and Research Methodology) of the American Educational Research

Association, 2(1).

Engelhard, G. (1991). Thorndike's essentials of a valid scale. *Rasch Measurement Transactions*, 5(3), 170.

Engelhard, G. (1991). Thorndike and Wood. *Rasch Measurement Transactions*, 5(2), 146.

Engelhard, G. (1991). Research traditions and history of measurement. *Rasch Measurement Transactions*, 4(4), 26.

Engelhard, G. (1990). Gender differences in performance on mathematics items: Evidence from the U.S. and Thailand. *Contemporary Educational Psychology*, 15, 13-26.

Tocci, C.M., & Engelhard, G. (1990). Mathematics achievement, parental support and gender differences in attitudes toward mathematics of American and Thai adolescents. *Georgia Educational Researcher*, 7, 16-18. [Summary of the 1989 Distinguished Paper awarded by Georgia Educational Research Association]

Engelhard, G., Hansche, L., & Rutledge, K.E. (1990). Accuracy of bias review judges in identifying differential item functioning on teacher certification tests. *Applied Measurement in Education*, 3(4), 347-360.

Engelhard, G., Anderson, D., & Gabrielson, S. (1990). An empirical comparison of Mantel-Haenszel and Rasch procedures for studying differential item functioning on teacher certification tests. *Journal of Research and Development in Education*, 23(2), 172-179.

Engelhard, G. (1990). Math anxiety, mother's education, and the mathematics performance of adolescent boys and girls: Evidence from the U.S. and Thailand. *Journal of Psychology*, 124(3), 289-298.

Monsaas, J.A., & Engelhard, G. (1990). Home environment and the competitiveness of highly accomplished individuals in four talent fields. *Developmental Psychology*, 26, 264-268.

Engelhard, G. (1990). History and philosophy of measurement. *Rasch Measurement Transactions*, 4(3), 118.

Engelhard, G. (1990). History of measurement. *Rasch Measurement Transactions*, 4(2), 110.

Engelhard, G., & Monsaas, J.A. (1989). Academic performance, gender and the cooperative attitudes of third, fifth and seventh grade students. *Journal of Research and Development in Education*, 22, 13-17.

Engelhard, G., & Monsaas, J.A. (1988). Grade level, gender and school-related curiosity in urban elementary schools. *Journal of Educational Research*, 82, 22-26.

Engelhard, G. (1988). Review of *Rasch Models for Measurement* by David Andrich. *Applied Psychological Measurement*, 12, 435-436.

Engelhard, G. (1986). A simulation study of computerized adaptive testing with a misspecified measurement model. *Proceedings of the American Statistical Association, Section on Survey Research Methods*. Washington, D.C.: American Statistical Association.

Engelhard, G. (1986). The discovery of educational goals and outcomes: A view of the latent curriculum of schooling. *Dissertation Abstracts International*, 46, No. 8, 2176-A.

Muraki, E., & Engelhard, G. (1985). Full-information item factor analysis: Applications of EAP scores. *Applied Psychological Measurement*, 9, 417-430.

Engelhard, G. (1984). Thorndike, Thurstone and Rasch: A comparison of their methods of scaling psychological and educational tests. *Applied Psychological Measurement*, 8, 21-38.

Engelhard, G., & Osberg, D.W. (1983). Constructing a test network with a Rasch measurement model. *Applied Psychological Measurement*, 7, 283-294.

Easton, J.Q., & Engelhard, G. (1982). A longitudinal record of elementary school absence and its relationship to reading achievement. *Journal of Educational Research*, 75, 269-274.

Guskey, T.R., Engelhard, G., Guida, F., Tuttle, K., & Roth, R. (1978). *Report on the pilot project to develop mastery learning courses for the Chicago Public Schools*. Chicago: Chicago Board of Education, Center for Urban Education.

Engelhard, G., & Monsaas, J.A. (1977). (Eds.). *Research strategies for selected educational problems*. Chicago: The University of Chicago, Department of Education.

Engelhard, G., Burke, A.J., & Yildiran, G. (1977). A social implication of Mastery Learning: Mastery Learning and the selection function of schools. In G. Engelhard & J.A. Monsaas (Eds.), *Research strategies for selected educational problems*. Chicago: The University of Chicago, Department of Education.

COMMISSIONED PAPERS, REPORTS AND MONOGRAPHS

Engelhard, G., Kobrin, J., Wind, S.A., & Chajewski, M. (2014). *Differential item and person functioning in large-scale writing assessments within the context of the SAT Reasoning Test*. College Board Research Report.

Engelhard, G., & Wind, S.A. (2013). *Rating Quality Studies using Rasch Measurement Theory*. College Board Research Report 2013-3.

Engelhard, G., Haertel, E., & Poggio, J. (2009). Report of technical review expert panel: Review of Oklahoma's modified alternate assessment program. Report commissioned by Oklahoma Dept. of Education (SRI International).

Randall, J. & Engelhard, G. (2008). *Using confirmatory factor analysis and Rasch Measurement Theory to assess measurement invariance in a high stakes reading assessment* (CEA Rep. No. 671). Amherst: University of Massachusetts, Center for Educational Assessment.

Randall, J. & Engelhard, G. (2008). *Examining teacher grades using Rasch Measurement Theory* (CEA Rep. No. 672). Amherst: University of Massachusetts, Center for Educational Assessment.

Engelhard, G., Fincher, M., & Domaleski, C. (2007) *Examining the Reading and Mathematics Performance of Students with Disabilities under Modified Conditions: The Georgia Department of Education Modification Research Study*. Report commissioned by the Georgia Department of Education and the Georgia Assessment Center.

Cizek, G.J., Engelhard, G., & Moody, M. (2005). *Review of Delaware student testing program assessments and performance standards*. Report commissioned by the Honorable Ruth Minner, Governor of Delaware in response to Executive Order 54.

Wolfe, E.W., Myford, C., & Engelhard, G. (2004). *Monitoring reader performance and DRIFT in the AP English Literature and Composition Exam using benchmark essays*. Report commissioned by the Advanced Placement Program of the College Board.

Engelhard, G., & Myford, C. (2002). *Monitoring reader performance in the Advanced Placement Program with a many-faceted Rasch model*. Report commissioned by the Advanced Placement Program of the College Board.

Engelhard, G., Myford, C., & Cline, F. (2000). *Investigating assessor effects in National Board for Professional Teaching Standards Assessments for Early Childhood/Generalist and Middle Childhood/Generalist Certification*. Report commissioned by the National Board for Professional Teaching Standards (NBPTS).

Tanzer, J., Elliott, J., Engelhard, G., Schrag, J., & Vogel, S. (2000). *Report of the Blue Ribbon Panel: Students with Learning Disabilities and the Oregon Statewide Assessment System*. Report commissioned by the Oregon Department of Education.

Engelhard, G. (1998). *Examination of psychometric issues raised by the Chicago Public*

Schools regarding norms and equating of the Iowa Tests. Report commissioned by The Riverside Publishing Company (Itasca, IL).

Engelhard, G. (1997 to present). Annual reports on the psychometric quality of the certification exams for nurse practitioners in areas related to women's health care. Reports commissioned by the National Certification Corporation (Chicago, IL).

Engelhard, G. (1997). *Exploratory analyses of gender differences in salary by pay grade at Emory University.* Report commissioned by the Staff Concerns Committee of the President's Commission on the Status of Women at Emory University.

Engelhard, G. (1994). *Rater consistency across time and domain in the assessment of written composition.* Report commissioned by the Georgia Department of Education.

Engelhard, G. (1994). *Examining the influences of rater adjustments on pass rates for the Georgia High School Graduation Writing Test.* Report commissioned by the Georgia Department of Education.

Engelhard, G. (1994). *Maintaining the ongoing psychometric quality of the National Board for Professional Teaching Standards (NBPTS) Assessments of highly accomplished teachers.* Report commissioned by the NBPTS and the University of North Carolina.

Loyd, B., Engelhard, G., & Crocker, L. (1993). *Equity, equivalence, and equating: Fundamental issues and proposed strategies for the National Board of Professional Teaching Standards (NBPTS).* Report commissioned by the NBPTS and the University of North Carolina.

Engelhard, G. (1992). *The assessment of written composition in Georgia.* Report commissioned by Test Scoring and Reporting Service of the Educational Research Laboratory, University of Georgia.

Engelhard, G. (1991). *Statistical issues in the measurement of writing performance with the Georgia Basic Skills Writing Test.* Report commissioned by Test Scoring and Reporting Service of the Educational Research Laboratory, University of Georgia.

PAPERS AND PRESENTATIONS GIVEN AT INTERNATIONAL, NATIONAL, REGIONAL, AND STATE CONFERENCES

Walker, A.A., & Engelhard, G. (August 2015). *Using Person Fit and Person Response Functions to Examine the Validity of Person Scores In Computer Adaptive Tests.* Paper to be presented to the Pacific Rim Objective Measurement Society (PROMS) in Fukuoka, Japan.

Behizadeh, N., & Engelhard, G. (August 2015). *Examining the Psychometric Quality of a Modified Perceived Authenticity in Writing Scale with Rasch Measurement Theory.* Paper to be presented to the Pacific Rim Objective Measurement Society (PROMS) in Fukuoka, Japan.

Ruiz-Menjivar, J., Engelhard, G., & Grable, J. (2015) *Evaluating the Psychometric Quality of the Grable and Lytton's Financial Risk Tolerance Scale Using Rasch Measurement*

Theory. Poster Session at the 2015 CFP Board Registered Programs Conference in Washington, D.C.

Wang, J., & Engelhard, G. (August 2015). *A Hyperbolic Cosine Unfolding Model for Evaluating Rater Accuracy in Writing Assessments*. Paper to be presented to the Pacific Rim Objective Measurement Society (PROMS) in Fukuoka, Japan.

Wind S. A., Engelhard, Jr., G., & Wesolowski, B. C., (2015, February). *Exploring the effects of rater linking designs and rater fit on achievement estimates within the context of music performance assessments*. Paper to be presented annual meeting of the American Educational Research Association Meeting in Chicago.

Wesolowski, B. C., Wind S. A., & Engelhard, Jr., G. (2015, February). *Evaluating Rater Quality in Music Performance Assessment*. To be presented at the Tenth Biennial Suncoast Music Education Research Symposium (SMERS X). Tampa, FL.

Wesolowski, B. C., Wind, S. A., & Engelhard, Jr., G. (2015, February) *Rater Analyses in Music Performance Assessment: Application of the Many Facet Rasch Model*. To be presented at the 5th International Symposium on Assessment in Music Education (ISAME), Connecting Practice, Measurement, and Evaluation. Fredricksburg, VA.

Engelhard, G. (2015). *Invariant measurement with raters and rating scales*. Invited keynote address to Rasch Measurement Special Interest Group to be presented at the annual meeting of the American Educational Research Association Meeting in Chicago.

Engelhard, G. (2015). *An Overview of Model-Based Indices for Monitoring Rating Quality*. Paper to be presented at the annual meeting of the National Council on Measurement in Education in Chicago.

Wind, S.A., Engelhard, G., (2015). *Exploring the Quality of Expert Ratings Using Mokken Scale Analysis*. Paper to be presented at the annual meeting of the National Council on Measurement in Education in Chicago.

Patterson, B.F., Wind, S.A., & Engelhard, G. (2015). *Incorporating Expert Ratings into Rater Monitoring via Signal Detection Theory*. Paper to be presented at the annual meeting of the National Council on Measurement in Education in Chicago.

Engelhard, G. (2014). *Invariant measurement with raters*. Paper presented at the International Test Commission meeting in San Sebastian, Spain.

Wind, S.A., Wolfe, E. W., Engelhard, G. Jr., Foltz, P., & Rosenstein, M. (2014). *Influence of rater effects on the training of automated scoring engines*. Paper presented at the International Test Commission meeting in San Sebastian, Spain.

Engelhard, G. (2014). *Invariant Measurement with Raters and Rating Scales*. Invited keynote address in the Pacific Rim Objective Measurement Society meeting in Guangzhou, China.

Razynski, K., Engelhard, G., Cohen, A., & Lu, Z (L.). (2014). *Comparing the Effectiveness of Rater Training Methods Using Generalized Linear Mixed Models*. Paper presented at the International Meeting of the Psychometric Society in Madison, WI.

Wang, J., & Engelhard, G., (2014). *Everyone's Rasch Measurement Analyzer (Erma): An R Package for Rasch Measurement Models*. Paper presented at the International Objective Measurement Workshop in Philadelphia, PA.

Behizadeh, N., & Engelhard, G. (2014). *What is a valid writing assessment?*. Paper presented at the International Objective Measurement Workshop in Philadelphia, PA.

Walker, A.A., Royal, K., Hedgpeth, M., & Engelhard, G., Jr. (2014) *Exploring Aberrant Responses Using Person Fit and Person Response Functions*. Paper presented at the annual meeting of the American Educational Research Association, Philadelphia.

Wind, S. A., & Engelhard, G., Jr. (2014) *Monitoring Rating Quality in Writing Assessment using Mokken Scale Analysis*. Paper to be presented at the annual meeting of the American Educational Research Association, Philadelphia.

Wind, S.A., Kaliski, P., Morgan, D., & Engelhard, G. (2013). *An Illustration of the Many-facet Rasch Model for Evaluating Modified-Angoff Standard-setting Judgments Within the Context of the AP World History Examination*. Paper presented at the meeting of the Northeastern Educational Research Association, Connecticut.

Cheong, Y.F., Wind, S.A., Engelhard, G., & Kaliski, P. (2013). *An Examination of the Hierarchical Cross-classified Modeling Approach to Standard Setting Ratings Within the Context of the AP World History Examination*. Paper presented at the meeting of the Northeastern Educational Research Association, Connecticut.

Wind, S. A., & Engelhard, G., Jr. (2013). *Examining Rating Quality in Writing Assessment using Mokken Scale Analysis and Rasch Measurement Theory*. Paper presented at the meeting of the European Association of Psychological Assessment, San Sebastian, Spain.

Perkins A.F., and Engelhard, G. (2013). *Examining Erasures in a Large-Scale Assessment of Mathematics and Reading*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Wind, S. A., & Engelhard, G., Jr. (2013) *How invariant and accurate are domain ratings in writing assessment?* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Walker, E.R., Barmon, C., Engelhard, G., Sterk, C.E., DiIorio, C. Thompson, N. (2012). *Experiences of living with epilepsy: An insiders' perspective from adults with epilepsy and their support persons* [abstract]. In Proceedings of the 140th Annual Meeting of the American Public Health Association. October 29, San Francisco, CA. Abstract number: 268544.

Walker, E.R., Barmon, C., Engelhard, G., Sterk, C.E., DiIorio, C. Thompson, N. (2012). *Qualitative analysis of support for self-management behaviors: Perspectives of adults with*

epilepsy and their support persons [abstract]. In Proceedings of the 140th Annual Meeting of the American Public Health Association. October 29, San Francisco, CA. Abstract number: 268595.

Tan, S., Chew, M. L., Wind, S. A., & Engelhard, G., Jr. (2012, August). *Evaluating the quality of ratings in writing assessment: Indices from Rasch measurement theory*. Paper presented at the Pacific Rim Objective Measurement Seminar, Shanghai.

Chang, M.L., & Engelhard, G. (2012). *Using the many-facet Rasch model to evaluate the psychometric quality of teacher sense of efficacy scales*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, CA.

Engelhard, G., & Wind, S.A. (2012). *Evaluating rater-mediated assessments*. Paper presented at the International Test Commission meeting in Amsterdam, July 3-5, 2012.

Engelhard, G., Kobrin, J., Wind, S.A., & Chajewski, M. (2012). *Differential item and person functioning in large-scale writing assessments within the context of the SAT Reasoning Test*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, CA.

Engelhard, G., & Wind, S.A. (2012). *Rating Quality Studies using Rasch Measurement Theory*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, CA.

Hendrickson, A., Engelhard, G., & Wind, S.A. (2012). *Comparative analyses of Generalizability Theory and the Many-Facet Rasch Model*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, CA.

Cheong, Y.F., Engelhard, G., & Wind, S.A. (2012). *Hierarchical Rasch models for rater-mediated assessments*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, CA.

Kaliski, P., Engelhard, G., Morgan, D., Reshetar, R., Wind, S.A., & Plake, B. (2012). *Using the Many-Facet Rasch Model to Evaluate Standard-Setting Judgments: Setting performance standards for Advanced Placement examinations*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, CA

Engelhard, G. (2012). *Rater-invariant assessments in the human sciences*. Paper presented at the Rasch Measurement Conference, University of Western Australia, Perth, Western Australia.

Wind, S.A., & Engelhard, G. (2012). *Examining Gender Differences in Writing Assessment using Rating Quality Indices*. Paper presented at International Objective Measurement Workshop (Rasch measurement) at the University of British Columbia, Vancouver, Canada.

Hendrickson, A., & Engelhard, G. (2011, July). *Evaluating the Quality of Rater-Mediated Assessments with a Multi-Method Approach*. Paper presented at the International Meeting of the Psychometric Society (IMPS) in Hong Kong.

Engelhard, G., & Behizadeh, N. (2011, July). *Exploring the Accuracy of Writing Self-Efficacy Using Rasch Measurement and Qualitative Methods*. Paper presented at the International Meeting of the Psychometric Society (IMPS) in Hong Kong.

Engelhard, G. (2011, July). The quest for invariant rater-mediated measurement within the context of performance assessments. Invited keynote address presented at the Pacific Rim Objective Measurement Seminar (PROMS) in Singapore.

Wind, S.A., & Engelhard, G. (2011, July) *Evaluating the quality of ratings in writing assessment: Rater agreement, precision, and accuracy*. Paper presented at the Pacific Rim Objective Measurement Seminar (PROMS) in Singapore.

Gale, J., & Engelhard, G. (2011, July). *Rasch analyses and mixed methods research: An investigation of urban elementary teachers' beliefs about their science teaching context*. Paper presented at the Pacific Rim Objective Measurement Seminar (PROMS) in Singapore.

Engelhard, G., & Behizadeh, N. (2011). *Exploring the Accuracy of Writing Self-Efficacy Judgments of 8th Graders Using Rasch Measurement Theory and Qualitative Methods*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Cheong, Y.F., Randall, J., & Engelhard, G. (2011). *Item analysis by hierarchical and cross-classified models: An application to a statewide reading assessment program*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans.

Engelhard, G. (2010). *The quest for invariant measurement in the human sciences*. Invited address at Fordham University, Bronx, New York.

Engelhard, G. (2010). Invited lecture series at The College Board (NYC)

- *The quest for invariant measurement in the human sciences (September 2010)*
- *Conceptualizing Differential Item and Person Functioning as Model-Data Misfit Fit (October 2010)*
- *Evaluating the quality of rater-mediated assessment (December 2010)*

Engelhard, G. (2010). *Performance assessments: Evaluating the quality of rater-mediated assessments*. Invited address at Pacific Rim Objective Measurement Workshop at International Islamic University Malaysia in Kuala Lumpur.

Engelhard, G. (2010). *The Quest for Invariant Measurement*. Paper presented at the international conference on Rasch measurement in Copenhagen, Denmark.

Clark, K., & Engelhard, G. (2010). *Evaluating the Psychometric Quality of a Racial Attitudes Index with a Rasch Measurement Model*. Paper presented at the international conference on Rasch measurement in Copenhagen, Denmark.

Engelhard, G. (2010). *An overview of historical and philosophical perspectives on measurement*. Keynote symposium to be presented at the meeting of the International Objective Measurement Workshop at the University of Colorado, Boulder.

Behizadeh, N., & Engelhard, G. (2010). *Influences of measurement and writing theories on writing assessment: An historical perspective*. Paper presented at the meeting of the International Objective Measurement Workshop at the University of Colorado, Boulder.

Quaynor, L., & Engelhard, G. (2010). *Historical View of Measurement and Language Theories Within the Context of the Assessment the Language Proficiency of English Language Learner*. Paper presented at the meeting of the International Objective Measurement Workshop at the University of Colorado, Boulder.

Keneman, M.L., Piot, S., & Engelhard, G. (2010). *Perennial Problems in the Assessment of Communicative Competence in a Second Language: A Rasch Perspective*. Paper presented at the meeting of the International Objective Measurement Workshop at the University of Colorado, Boulder.

Perkins, A., & Engelhard, G. (April 2010). *Using differential person functioning to examine student performance on mathematics items*. Paper presented at the meeting of the International Objective Measurement Workshop at the University of Colorado, Boulder.

Perkins, A., Quaynor, L., & Engelhard, G. (April, 2010). *Crossing Person Response Functions: The Influences of Home Language, Gender, and Social Class on Mathematics Literacy in France, Germany, Hong Kong, and the United States*. Paper presented at the 2010 American Educational Research Association Annual Meeting.

Behizadeh, N., & Engelhard, G. (April 2010). *Historical View of the Influences of Measurement and Writing Theories on the Practice of Writing Assessment*. Paper presented at the 2010 American Educational Research Association Annual Meeting.

Caines, J., & Engelhard, G. (November 2009). *The Nexus between Measurement and Evaluation: A Process Evaluation of a Standard Setting Procedure using Rasch Measurement Theory and Evaluation Theory*. Paper presented at the annual meeting of the American Evaluation Association in Orlando, Florida.

Quaynor, L., Perkins, A., & Engelhard, G. (July 2009). *An analysis of differential item functioning and differential person functioning on the PISA mathematics test for France, Germany, Hong Kong, and the United States*. Paper presented at Pacific Rim Objective Measurement Symposium (PROMS). Hong Kong.

Caines, J., & Engelhard G. (July 2009). *An application of Rasch measurement in setting educational standards*. Paper presented at Pacific Rim Objective Measurement Symposium (PROMS). Hong Kong.

Engelhard, G. (July 2009). *The quest for invariant measurement in the social, behavioral and health sciences*. Invited address presented at Pacific Rim Objective Measurement Symposium (PROMS). Hong Kong.

Caines, J., & Engelhard, G. (April 2009). *Evaluating Body of Work Judgments of Standard-Setting Panelists*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Engelhard, G. (August 2008). *Historical view of the influences of theories of measurement and language proficiency on English language tests*. Keynote address presented at Pacific Rim Objective Measurement Symposium (PROMS). Tokyo.

Engelhard, G. (March 2008). *Historical view of theories of measurement and language proficiency within the context of second language testing*. Paper presented at the 2008 meeting of the International Objective Measurement Workshop. New York: New York University.

Caines, J., & Engelhard, G. (March 2008). *Using the Rating Scale Model to Examine Angoff Ratings of Standard-Setting Panelists*. Paper presented at the 2008 meeting of the International Objective Measurement Workshop. New York: New York University.

Sullivan, R., Caines, J., Tucker, C., & Engelhard, G. (March 2008). *Examining the bookmark ratings of standard-setting panelists: An approach based on the multifaceted Rasch measurement model*. Paper presented at the 2008 meeting of the International Objective Measurement Workshop. New York: New York University.

Engelhard, G. (December 2007). *Constructing Measures with the Many-Faceted Rasch Model*. Invited distinguished lecture series, Temple University Japan, Tokyo and Osaka.

Engelhard, G. (December 2007). *Fundamentals of Psychometric Theory*. Invited lecture to National Certification Corporation, Chicago, IL.

Engelhard, G. (July 2007). *Conceptualizing Differential Item and Examinee Functioning as Model-Data Misfit for Students with Disabilities*. Paper presented at the International Meeting of the Psychometric Society, Tokyo, Japan.

Engelhard, G. (April 2007). *Historical perspectives on invariant measurement: Thurstone, Rasch, Guttman, and Mokken*. Invited address, Rasch Special Interest Group, American Educational Research Association. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Thomas, J.R., & Engelhard, G. (April 2006). *Using Guttman's Facet Theory to Develop an Instrument that Examines the Grading Practices of Teachers*. Paper presented at the 13th International Objective Measurement Workshop at University of California, Berkeley.

Alviar, T.B., Usher, E., Thomas, J., & Engelhard, G. (November 2005). *Teaching topics in Civics Education: Functions of country and degree of teacher confidence*. Paper presented at the annual meeting of the Georgia Educational Research Association in Savannah, GA.

Alviar, T.B., Usher, E., Thomas, J., & Engelhard, G. (April 2006). *Teaching topics in Civics Education: Influences of degree, gender, and country on teacher confidence*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Alviar, T.B., Usher, E., Thomas, J., & Engelhard, G. (January 2006). *Teaching topics in Civics Education: A political socialization perspective*. Paper presented at the 4th Annual Civic Education Conference, Altamonte Springs, Florida.

Myford, C.M., Wolfe, E.W., Engelhard, G., & Manalo, J.R. (April 2005). *Monitoring rater performance and DRIFT in the Advanced Placement English Literature and Composition Examination using benchmark essays*. Paper presented at the annual meeting of the National Council on Measurement in Education (Montreal).

Engelhard, G. (July 2004). *Issues associated with implementing "No Child Left Behind" from the perspective of the technical advisor*. Invited paper presented at the annual meeting of the Council of Chief State School Officers in Boston, MA.

Engelhard, G. (December 2003). *Using Rasch measurement to evaluate the ratings of standard-setting judges*. Presentation at the Università degli Studi di Padova (University of Padua) in Padua, Italy.

Engelhard, G. (April 2003). *Thurstone, Guttman, and Rasch: A comparison of their perspectives on invariant measurement*. Paper presented at the Wright Conference in Chicago.

Engelhard, G. (July 2003). *Evaluating rater-mediated assessments*. Paper presented at the International Meeting of the Psychometric Society, Sardinia, Italy.

Engelhard, G., & Myford, C. (2002, April). *Monitoring reader performance in the Advanced Placement English Literature and Composition Program with a Many-Faceted Rasch Model*. Paper presented at the International Objective Measurement Workshop (IOMW11) in New Orleans.

Engelhard, G., & Myford, C. (2001, July). *Using Rasch measurement to examine reader severity on the Advanced Placement Exams in English Literature/Composition*. Paper presented at the International Meeting of the Psychometric Society, Osaka, Japan.

Engelhard, G., & Myford, C. (2001, April). *Monitoring reader performance in the Advanced Placement English Literature and Composition Program with a Many-Faceted Rasch Model*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Hambleton, R., & Engelhard, G. (2001, April). *Richard M. Jaeger's contributions to Standard Setting practices*. Paper presented at the annual meeting of the American Educational

Research Association, Seattle, WA.

Bazarova, S.I., & Engelhard, G. (2001, April). *Using Rasch Measurement Theory to explore cross-cultural DIF on mathematics items: Evidence from the United States and the Russian Federation*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Calvert, T., & Engelhard, G. (2000, April). *Gender differences on eighth grade mathematics items: A cross-cultural comparison of the United States and Spain*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Engelhard, G., Myford, C., & Cline, F. (2000, April). *Examining the influences of assessor severity on the certification of highly accomplished teachers*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Bazarova, S.I., & Engelhard, G. (2000, April). *Using Rasch Measurement Theory to explore translation DIF on mathematics items: Evidence from the United States and the Russian Federation*. Paper presented at the 10th International Objective Measurement Workshop, New Orleans, LA.

Garner, M., & Engelhard, G. (2000, April). *An eigenvector method for estimating item parameters of the dichotomous and polytomous Rasch models*. Paper presented at the 10th International Objective Measurement Workshop, New Orleans, LA.

Garner, M., & Engelhard, G. (2000, April). *Rasch measurement as an Analytic Hierarchy Process*. Paper presented at the 10th International Objective Measurement Workshop, New Orleans, LA.

Calvert, T., & Engelhard, G. (1999, October). *Gender differences on eighth grade mathematics items: A cross-cultural comparison of the United States and Spain*. Paper presented at the annual meeting of the Georgia Educational Research Association, Atlanta, GA.

Bazarova, S.I., & Engelhard, G. (1999, October). *Cross-cultural comparisons of the mathematics achievement of eighth grade students in the United States and the Russian Federation*. Paper presented at the annual meeting of the Georgia Educational Research Association, Atlanta, GA.

Karunungan, M.L.L., & Engelhard, G. (1999, October). *Mother's education and the mathematics performance of eighth grade girls and boys in Singapore and the United States*. Paper presented at the annual meeting of the Georgia Educational Research Association, Atlanta, GA.

Gyagenda, I., & Engelhard, G. (1999, April). *Using classical and modern measurement theories to explore rater, domain, and gender influences on student writing ability*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Engelhard, G., & Gordon, B. (1998, January). *New procedures for setting performance standards*. Paper presented at the International Conference on Latent Trait Theory: Rasch Measurement, University of Western Australia, Perth, Australia.

Gyagenda, I., & Engelhard, G. (1998, April). *Rater, domain, and gender influences on the assessment of quality of student writing using weighted and unweighted scoring*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Gyagenda, I., & Engelhard, G. (1998, April). *Applying the Rasch model to explore rater influences on the assessed quality of students' writing ability*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Garner, M., & Engelhard, G. (1998, April). *The application of graph theory to issues of connectivity in assessment networks*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Engelhard, G., & Gordon, B. (1997, July). *Setting and evaluating performance standards for high stakes writing assessments*. Paper presented at the 10th European Meeting of the Psychometric Society, Santiago de Compostela, Spain.

Monsaas, J.A., & Engelhard, G. (1997, July). *Using the many-facet Rasch model to examine gender differences in attitudes toward computers*. Paper presented at the 10th European Meeting of the Psychometric Society, Santiago de Compostela, Spain.

Engelhard, G., & Pajares, F. (1997, March). *Measuring the accuracy of self-efficacy judgments in mathematics*. Paper presented at the Ninth International Objective Measurement Workshop, The University of Chicago, Chicago, IL.

Monsaas, J.A., & Engelhard, G. (1997, March). *Gender differences in attitudes toward computers*. Paper presented at the Ninth International Objective Measurement Workshop, The University of Chicago, Chicago, IL.

Garner, M., & Engelhard, G. (1997, March). *The method of paired comparisons*. Paper presented at the Ninth International Objective Measurement Workshop, The University of Chicago, Chicago, IL.

Engelhard, G., & Gordon, B. (1997, March). *Setting and evaluating performance standards for high stakes writing assessments*. Paper presented at the Ninth International Objective Measurement Workshop, The University of Chicago, Chicago, IL.

Stone, G.E., & Engelhard, G. (1997, March). *Informing mastery through binomial trails: A refinement of Objective Standard Setting*. Paper presented at the Ninth International Objective Measurement Workshop, The University of Chicago, Chicago, IL.

Garner, M. L., & Engelhard, G. (1997, March). *Paired comparisons, graph theory, and Rasch measurement theory*. Paper presented at the annual meeting of the American Educational

Research Association, Chicago, IL.

Engelhard, G., & Gordon, B. (1996, October). *Setting and evaluating performance standards for high stakes writing assessments*. Paper presented at the annual meeting of the Georgia Educational Research Association, Atlanta, GA.

Garner, M.L., & Engelhard, G. (1996, October). *Analysis of an assessment network using graph theory*. Paper presented at the annual meeting of the Georgia Educational Research Association, Atlanta, GA.

Engelhard, G. (1996, April). *Models of judgment and Rasch measurement theory*. Paper presented at the annual meeting of the American Educational Research Association, New York, New York.

Engelhard, G., Davis, M., & Hansche, L. (1996, April). *Evaluating the accuracy of judgments obtained from item review committees*. Paper presented at the annual meeting of the American Educational Research Association, New York, New York.

Garner, M. L., & Engelhard, G. (1996, April). *Gender differences in performance on multiple-choice and constructed response mathematics items*. Paper presented at the annual meeting of the American Educational Research Association, New York, New York.

Monsaas, J.A., Matthews, M.W., & Engelhard, G. (1996, April). *Identifying rater errors in the analysis of text using the many-faceted Rasch model*. Paper presented at the annual meeting of the American Educational Research Association, New York, New York.

Gordon, B., & Engelhard, G. (1995, October). *How accurately can readers identify the gender of student writers?* Paper presented at the annual meeting of the Georgia Educational Research Association, Atlanta, GA.

Garner, M. L., & Engelhard, G. (1995, October). *Paired comparisons, graph theory, and Rasch measurement theory*. Paper presented at the annual meeting of the Georgia Educational Research Association, Atlanta, GA.

Engelhard, G. (1995, October). *Introduction to the many-faceted Rasch measurement model*. Pre-session presented at the annual meeting of the Georgia Educational Research Association, Atlanta, GA.

Engelhard, G., & Anderson, D.W. (1995, July). *Examining changes in the ratings of standard-setting judges*. Paper presented at the 9th European Meeting of the Psychometric Society, Leiden, The Netherlands.

Engelhard, G. (1995, April). *Models of judgment and educational assessment*. Paper presented at the Eighth International Objective Measurement Workshop at the University of California, Berkeley.

Engelhard, G., & Cramer, S. (1995, April). *Using Rasch Measurement to evaluate the*

ratings of standard-setting judges. Paper presented at the Eighth International Objective Measurement Workshop at the University of California, Berkeley.

Engelhard, G., & Anderson, D.W. (1995, April). *Examining changes in the ratings of standard-setting judges over time*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Engelhard, G. (1995, March). *Models of judgment and educational assessment*. Paper presented at the annual meeting of the Eastern Educational Research Association, Hilton Head, South Carolina.

Engelhard, G. (1994, November). *Overview of the many-faceted Rasch model*. Presentation to the Michigan Department of Education, Lansing, MI.

Engelhard, G. (1994, July). *Evaluating rater accuracy in performance assessments with a many-faceted Rasch model*. Paper presented at the 23rd International Congress of Applied Psychology, Madrid, Spain.

Engelhard, G., Gordon, B., & Curtin, D. (1994, April). *Constructing rater and writing task banks for the assessment of written composition*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Gabrielson, S.J., Gordon, B., & Engelhard, G. (1994, April). *The effects of topic choice on the quality of writing obtained in a statewide assessment*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Engelhard, G., Davis, M., & Hansche, L. (1993, December). *Evaluating the accuracy of judgments obtained from item review committees*. Paper presented at the Midwest Objective Measurement Seminar at the University of Chicago.

Engelhard, G. (1993, July). *Evaluating the accuracy of expert judgments of writing competence with benchmark compositions*. Paper presented at the European meeting of the Psychometric Society, Barcelona, Spain.

Monsaas, J.A., & Engelhard, G. (1993, July). *Using the Rasch model to examine changes in the home environment*. Paper presented at the European meeting of the Psychometric Society, Barcelona, Spain.

Lloyd, B., Engelhard, G., & Crocker, L. (1993, April). *Equity, equivalence, and equating: Fundamental issues and proposed strategies for the National Board of Professional Teaching Standards*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Gordon, G., Engelhard, G., Gabrielson, S., & Bernknopf, S. (1993, April). *When you need to equate performance assessments . . . carefully consider your options*. Paper presented at the annual meeting of the American Educational Research Association.

Monsaas, J.A., & Engelhard, G. (1993, February). *Examining changes in the home*

environment with Rasch measurement models. Paper presented at the annual meeting of the Eastern Educational Research Association, Clearwater, FL.

Engelhard, G. (1992, June). *The assessment of written composition*. Paper presented at the first meeting of the European Conference on Educational Research, Enschede, The Netherlands.

Engelhard, G., Gordon, B., Siddle Walker, E.V., & Gabrielson, S. (1992, March). *Influences of writing tasks and gender on the quality of writing for black and white students*. Paper presented at the annual meeting of the Eastern Educational Research Association, Hilton Head Island, NC. [Received the Distinguished Research Award].

Engelhard, G., & Cramer, S.E. (1992, April). *The influences of item characteristics on judge consistency within the context of standard-setting committees*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Engelhard, G. (1992, April). *Constructing a rater bank for the assessment of written composition*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Engelhard, G. (1991, December). *Issues in the assessment of written composition*. Paper presented at the annual Midwest Objective Measurement Seminar at the University of Chicago.

Engelhard, G. (1991, April). *The measurement of writing ability with a many-faceted Rasch model*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Engelhard, G., Gordon, B., & Gabrielson, G. (1991, April). *Writing tasks and the quality of student writing: Evidence from a statewide assessment of writing*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Engelhard, G. (1991, April). *Historical views of the concept of invariance and measurement theory in the behavioral sciences*. Paper presented at the Sixth International Objective Measurement Workshop, The University of Chicago, Chicago, IL.

Monsaas, J.A., & Engelhard, G. (1991, April). *Predicting teachers' cheating on standardized tests*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL. [Presentation of the 1990 GERA Distinguished Paper]

[Thirty papers and presentations between April 1980 and March 1991]

ORGANIZATIONAL MEMBERSHIPS

American Educational Research Association (Fellow)

National Council on Measurement in Education

The Psychometric Society

International Test Commission

PROFESSIONAL SERVICES: International, National and State

Pacific Rim Objective Measurement Society

Board member (2012 to present)

American Educational Research Association

Program Chair, Division D (Measurement and Research Methodology), American Educational Research Association (1992 annual conference in San Francisco). Responsibilities included coordination of program activities related to three sections (Applied measurement, theoretical measurement, and research methods) within Division D.

Section Chair, Section 2 (Theoretical measurement) of Division D (Measurement and Research Methodology), American Educational Research Association (1991 annual conference in Chicago). Responsibilities included coordination of the peer-review process and selection of papers for presentation at the annual meeting.

Co-Chair, Graduate Student Seminar Committee, American Educational Research Association, (1997-98). Responsibilities included organizing a pre-session to provide graduate students with the opportunity to participate in round table discussions with distinguished scholars regarding current topics in research methodology, statistics, and measurement.

Chair, Rasch Measurement Special Interest Group, American Educational Research Association (1994-96).

Member, E.F. Lindquist Award Committee, American Educational Research Association (1991-1993). The purpose of this award is to recognize outstanding lifetime contributions in educational measurement.

Proposal reviewer, American Educational Research Association, Division D, Measurement and Research Methodology, Sections on Applied and Theoretical Measurement (1986-2000).

Proposal reviewer, American Educational Research Association, Special Interest Group on Rasch Measurement (1987, 1993, 1995).

Proposal reviewer, American Educational Research Association, Special Interest Group on the National Assessment of Educational Progress (1985, 1986).

National Council on Measurement in Education

Chair, NCME Website Taskforce. The purpose of the taskforce is to review the current NCME website and recommend improvements in the website (2006-2007).

Member, Brenda Loyd Dissertation Award Committee, National Council on Measurement in Education (2001-2002)

Editor, Instructional Materials for Educational Measurement Series (ITEMS), National Council of Measurement in Education (1994 to 1997)

Member, Instructional Materials for Educational Measurement Series (ITEMS), National Council of Measurement in Education (1997 to present)

Member, NCME Award for Career Contributions to NCME Committee, National Council on Measurement in Education, (1997-1998).

Member, Annual Award Committee, National Council on Measurement in Education, (1992-1994).

Proposal reviewer, National Council on Measurement in Education, (1990-2000).

Georgia Educational Research Association

President Elect, President, Past President, Georgia Educational Research Association, 1991-92, 1992-93, 1993-94.

Member, Nomination Committee, Georgia Educational Research Association, (Fall 1987).

Editorial Boards

Associate Editor (2009 to 2014), *Applied Measurement in Education*. Appointment as co-editor pending.

Member of Editorial Board (2004 to present), *Educational and Psychological Measurement*.

Member of Editorial Review Board (1996 to present), *Applied Measurement in Education*.

Founding Member of Editorial Board (2001 to present), *Measurement: Interdisciplinary Research and Perspectives*.

Founding Member of Editorial Board (2000 to present), *Journal of Applied Measurement*.

Founding Member of Educational/Psychological Editorial Board (1997 to 2001), *Journal of Outcome Measurement*.

Member of Editorial Advisory Board (1997 to 2001), *Educational Measurement: Issues and Practice*.

Member of Board of Consultants (1997 to 2000), *Journal of Research and Development in Education*

Manuscript reviewer

American Educational Research Journal, Review of Educational Research, Journal of Educational Measurement, Educational Assessment, Journal of Educational Statistics, Georgia

Educational Researcher, Child Development, Educational Measurement: Issues and Practice, Research in the Teaching of English, Applied Measurement in Education, Applied Psychological Measurement, Applied Psychology: An International Review Journal of Applied Social Psychology, Language Testing, British Journal of Mathematical and Statistical Psychology, Journal of Educational Psychology, Journal of Occupational and Organizational Psychology, International Journal of Testing

International Conference on Measurement

Organizer with J. Monsaas (April, 1993), Seventh International Objective Measurement Workshop at Emory University, Atlanta, GA. The purpose of this workshop was to bring together international scholars conducting research on objective measurement in the social and behavioral sciences.

National and Statewide Technical Advisory Committees

Member, National Technical Advisory Committee on Dynamic Learning Maps (assessing students with disabilities). (January 2015 to present)

Member, Pennsylvania Technical Advisory Committee on teacher evaluation and certification. (July 2014 to present)

Member, Technical Advisory Committee, Georgia Department of Education. Responsibilities include advising the GDE on measurement and statistical issues related to student and teacher assessment (1986 to present).

Member, Technical Advisory Committee, Louisiana Department of Education. Responsibilities include advising the LDOE on measurement and statistical issues related to student assessment (2008 to present).

Member, Technical Advisory Committee on Standard Setting, ACT and National Assessment of Educational Progress (NAEP). Responsibilities include advising ACT and NAEP on standard setting procedures for the new Economics Test (2006 to 2007).

Member, Technical Advisory Committee, Pennsylvania Department of Education. Responsibilities include advising the PDOE on measurement and statistical issues related to student assessment (2008 to present).

Member, Technical Advisory Committee, New Jersey Department of Education. Responsibilities include advising the NJDOE on measurement and statistical issues related to student assessment (2001 to present).

Member, Technical Advisory Committee, Ohio Department of Education. Responsibilities include advising the ODE on measurement and statistical issues related to student assessment (1996 to present).

Member, Technical Advisory Committee, Michigan Department of Education. Responsibilities

include advising the MDE on measurement and statistical issues related to student assessment (1997 to present).

Member, Technical Advisory Committee, Washington Department of Education. Responsibilities include advising the WDOE on measurement and statistical issues related to student assessment (2001 to present).

Member, Technical Advisory Committee, Mississippi Department of Education. Responsibilities include advising the MDE on measurement and statistical issues related to student and teacher assessment (2013 to present).

Chair, Technical Advisory Committee, Gwinnett County Schools, Georgia. Responsibilities include advising the Gwinnett County schools on measurement and statistical issues related to student assessment (Member: 2001 to 2013, Chair: 2002-2003, 2005-2006).

Member, Massachusetts Tests for Educator Licensure (MTEL) Technical Advisory Committee. (May 2009 to 2010)

Member, Technical Advisory Panel on Uniform National Rules for NAEP Testing of Students with Disabilities (June 2009)

Member, Technical Advisory Committee on Standard Setting, ACT and National Assessment of Educational Progress (NAEP). Responsibilities include advising ACT and NAEP on standard setting procedures for the Science Assessment (2008 to 2010).

Member, Advisory Panel, Measurement Issues and the Washington State Writing Assessment Program, Washington State Department of Education. (Seattle, WA, October 2000 to 2002).

Member, Technical Advisory Panel, State-Developed Alternative Assessment (SDAA), Texas State Department of Education (Austin, TX, November 2000)

Member, Blue Ribbon Panel on Test Adaptations for Children with Learning Disabilities. Portland, Oregon (December 1999 to 2001).

Member, Measurement Research Advisory Panel (MRAP) for National Board of Professional Teaching Standards (May 1997 to 2003).

Member, Measurement Panel, Georgia Department of Education. Responsibilities include advising the GDE on measurement and statistical issues related to the development of new statewide student assessment program: Criterion Referenced Competency Tests. These tests are designed to measure Georgia's revised Quality Core Curriculum (1997 to 2005).

Member, Technical Design Team, Oregon Department of Education. Responsibilities include advising the ODE on measurement and statistical issues related to student assessment (1998 to 2003).

Member, National Advisory Committee for Performance Assessments for Colleges and Technical Schools (PACTS) project. Responsibilities include advising PACTS on measurement

and statistical issues in the development of alternative assessment system for admission to post-secondary education in Georgia (1999 to 2003).

Member, Georgia Kindergarten Assessment Panel, Georgia Department of Education. Responsibilities include serving as technical advisor on measurement and statistical issues related to assessment of first grade readiness of kindergarten children (1997 to 1998).

Member, Department Technical Advisory Board for the Georgia Comprehensive Evaluation System. Georgia Department of Education. Responsibilities included advising the GDE on measurement and statistical issues related to the evaluation of Georgia schools (1989 to 1991).

Member, Writing Steering Committee. Georgia Department of Education. Responsibilities include advising the GDE on measurement and statistical issues related to the statewide assessment of written composition (1991 to 1995).

Short Biographical Statement

Professor George Engelhard, Jr., Ph.D. joined the faculty at The University of Georgia in the fall of 2013. He is professor emeritus at Emory University (1985 to 2013). Professor Engelhard received his Ph.D. in 1985 from The University of Chicago (MESA Program -- measurement, evaluation, and statistical analysis). His research has focused on a variety of topics including: invariant measurement; the role of assessment in improving educational processes; Rasch measurement theory; rater-mediated assessments; history of measurement theory; assessment of written composition; development of affective human characteristics; teacher assessment; and certification testing. Among his current research are projects on the following: invariant measurement; models of judgment for educational assessments; examination of rater effects in a variety of contexts; and standard setting on high-stakes performance assessments. Professor Engelhard is co-editor of four books, and the author or co-author of over 150 journal articles, book chapters, and monographs. He serves on several national technical advisory committees on educational measurement and policy including Georgia, Louisiana, Michigan, Mississippi, New Jersey, Ohio, Pennsylvania, and Washington. He has served as President of the Georgia Educational Research Association (1992-93). He has received several awards and fellowships including a National Academy of Education/Spencer Foundation Fellowship Award, a Lilly Post-Doctoral Teaching Award, and Sears Writing across the Curriculum Project Award. He is a fellow of the American Educational Research Association. His latest book was published in 2013: *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences* (New York: Routledge). He is currently under contract to write a new book entitled: *Invariant measurement with raters: Classical and modern measurement theories for rating scales* (New York: Routledge).

Gerunda B. Hughes, Ph.D.

Director, Office of Institutional Assessment, Howard University



ghughes@howard.edu

Dr. Gerunda Burke Hughes is currently the Director of the Office of Institutional Assessment and Evaluation in the Office of the Provost and Chief Academic Officer at Howard University and Professor in the [Department of Curriculum and Instruction](#) in the [School of Education](#). In her faculty position, she served as Coordinator of Secondary Education Programs and taught courses in mathematics methods, assessment and measurement, and research methodology. She has also taught courses in the Department of Mathematics and the Center for Academic Reinforcement (CAR). Dr. Hughes has a B.S. in Mathematics from the University of Rhode Island; a M.A. in Mathematics from the University of Maryland, College Park; and a Ph.D. in Educational Psychology from Howard University.

In addition to teaching, Dr. Hughes has been involved in various professional activities related to research and assessment. She served as a Principal Investigator of the "Classroom Assessment Project" (CAP) at Howard University's Center for Research on the Education of Students Placed at Risk (CRESPAR) and as Principal Investigator or Co-Principal Investigator of several grants funded by the National Science Foundation and the United States Department of Education. She has also served as a mentor to Howard University students in the Ronald E. McNair Post-Baccalaureate Achievement Program, the Rockefeller Brother's Fund and the Howard University Science, Engineering and Mathematics Program.

She served as Co-Editor-in-Chief of the esteemed Journal of Negro Education, as Associate Editor of the highly acclaimed AERA journal, Review of Educational Research, and on the editorial boards of the American Educational Research Journal and the Mathematics Teaching-Research Journal. Dr. Hughes has presented numerous papers and published several articles and book chapters on issues related to educational assessment and evaluation.

ROBERT LEE

PROFESSIONAL EXPERIENCE

2005 to Present Dept. of Elementary and Secondary Ed. Malden MA
Chief Analyst; Acting PARCC Coordinator

- Manager of the team of analysts responsible for producing all MCAS, NAEP and MEPA reports
- Chair of the PARCC Ad hoc committee on score reporting; member PARCC Field Test Technical Team; Ad Hoc Committee on Growth, Research and Development and Standard Setting Advisory Committee
- Developer of the state's growth model: Student Growth Percentiles
- Lead technical advisor on the MCAS and WIDA-ACCESS for ELLs contracts responsible for reviewing IRT and test equating work
- Lead efforts to reform the test equating process to allow for greater accuracy in MCAS reporting
- Developed numerous innovative ways to manage transitions as the state moved from 4, 8 and 10th grade testing to 3-8 and 10th grade testing

2001-2005 Department of Education Malden MA
Analyst/Database Manager

- Designed the Massachusetts Adequate Yearly Progress system, which pioneered the use of an index system instead of percent proficient
- Integrated MCAS data with SIMS data to establish the databases used for the state accountability and Competency Determination systems
- Designed Annual Measurable Achievement Objectives for programs serving English language learners
- Lead the team of analysts providing regular updates to the commissioner and board of education on the effects of the new Competency Determination policy

1988-1999 Patriot Ledger/Baltimore Sun Quincy MA/Baltimore MD
Reporter

- Daily beat reporter covering crime, politics and local government
- Author of 2,500 daily news stories
- AP award winner for spot new coverage
- Covered McDuffy vs. Robertson case and the subsequent passage of the Education Reform Act of 1994

EDUCATION

2008 Completed coursework in IRT and test equating taught by Ron Hambleton at UMASS and designed for government employees.

1999-2000 Harvard Graduate School of Education Cambridge MA
EdM. Research Methods and Statistics

- Studied Hierarchical Linear Modeling, Regression, questionnaire design
- Studied the history of education reform under Prof. Paul Reville
- Interned at the Department of Education

1998-1999 UMASS -Boston Boston MA
30 credits study in Teacher Education program

- Coursework focused on constructivism in secondary science education
- Worked as a full-time substitute and Alternative Education teacher for Canton Public Schools

1985-1989 Johns Hopkins University Baltimore MD
BA. The Writing Seminars

- Journalism and public policy focus, graduated with departmental honors
- 30 credits of Biology coursework
- Captain of the 1989 Division III national champion fencing team
- Completed four internships at the Baltimore Sun and Baltimore Magazine

ADDITIONAL PROFESSIONAL ACTIVITIES

Advisor to the Georgia Department of Education's Race to the Top Growth Model Committee

Staff liaison to the Massachusetts Educator Evaluation Task Force

Member since 2010 of the Missouri Growth Technical Advisory Committee

Member since 2006 of the National Blue Ribbon Advisory Committee responsible for choosing Blue Ribbon Schools for the Department of Education

Member since 2011 of the Mississippi Assessment Technical Advisory Committee

Professional Vitae
Annotated Work History

Prepared for:
Superintendent Dr. Carey Wright

Professional Vitae

JOHN PAUL BEAUDOIN

Quality Summary

My skill-set combines both formal and practical experience in the areas of assessment, accountability, and educator effectiveness. Having served as a classroom teacher, district supervisor, and state agent, I combine my technical expertise in psychometrics, statistics, and research with a practitioner's lens to provide understandable and actionable solutions. My experience and training as a United State Department of Education Peer Reviewer and as a private-sector vendor have created a national network of professionals throughout the education community. This network creates additional mechanisms to seek and share practices being used throughout the country. In numerous cases, these contacts have assisted state and local leaders in making decisions that inform teaching and learning. My decision-making approach goes beyond simple technical solutions, but integrates stakeholder perspectives, end-user needs, and multiple communication approaches. My focus on self-improvement and integrity provides the foundation for my leadership style.

Mississippi Department of Education (Jackson, MS)	<ul style="list-style-type: none">• 2015-Reviewed and monitored the standard-setting for the 3rd grade, statewide assessment (MKAS²); Facilitated the OTSS and Accountability office's integration of the programming used to create school and district accountability results; Developed and implemented the Excellence for All (E4A) assessment quality and accountability review process; Created the Student Learning Objective (SLO) training using <i>Homeroom</i>; Conducted the SLO training pilot at four (4) sites across Mississippi; Provided large-scale (web-based) training solutions for SLO and SSO; Developed and presented the graduation index used to combine student grades and SATP2 results.• 2014-Evaluated via auditing, the MDE's newly developed school accountability system. Created procedural guidelines and the data validation indicators used to ensure credible results. Co-developed Mississippi's Teacher Evaluation System (MTES) analytical model used to combine performance results, drafted the technical manual, and provided feedback regarding the Student Growth Percentile (SGP) metric. Presented MTES analytical model to stakeholder groups. Reviewed the business rules for school and district accountability based upon the recommendations of the Accountability Task Force.• 2013-Served as an external member of the state's Accountability Committee. Developed and implemented a customized process to create performance levels and associated "cut scores" for the statewide accountability system.
--	--



Education

Louisiana State University, Ed.D., Administration: Economics of Education

University of Louisiana at Lafayette, M.Ed., School Administration

Louisiana State University, BS, Special Education

Employment History

2004-current: Research in Action, Inc. (CEO)

2002-2004: St. Tammany Parish School District (Supervisor)

1998-2002: Louisiana Department of Education (Director: Accountability)

1991-1998: Special Education Teacher (Long Beach, CA; Lafayette, LA, Livingston, LA)

2001-current: Army National Guard: MAJ (Officer Candidate School, Camp Beauregard, LA)

1983-1988: United States Air Force: SGT (West Germany)

Work History

Consultation

<p>United States Department of Education (Washington, DC)</p>	<ul style="list-style-type: none"> • 2009-Developed the peer review process necessary to evaluate each state's graduation rate. Trained lead facilitators and then facilitated the three day event as the chairperson. Provided summary analytics to USED Title I director and his staff. • 2005-Provided technical assistance to the Wyoming Department of Education (WDE) regarding the evidence needed for the agency's alternate assessment. Trained the Commonwealth of Puerto Rico staff in preparing evidence for their standards and assessment review (initial). Evaluated alternative AYP components for the Miccosukee Tribe, Oceti Sakowin Education Consortium (OSEC), and the Navajo Nation. Trained selected tribal representatives on the requirements associated with each principle outlined within the <i>USED's Accountability Workbook</i>.
<p>US Department of the Interior- Bureau of Indian Education (Washington, DC)</p>	<ul style="list-style-type: none"> • 2012- Designed the technical elements of the BIE's ESEA Flexibility Request, including each metric associated with the Accountability Index (AI). Trained selected BIE staff in the accountability redesign elements and the assessment requirements necessary to support growth model data. • 2010- Selected by the BIE as the technical assistance provider for tribal entities throughout the nation who are seeking to develop and implement an alternate AYP definition. Provided on-site and remote services in developing Accountability Workbooks in a manner that meets the cultural needs of tribal communities. . Developed procedures for creating alternate AYP definitions consistent with the requirements of 25 C.F.R. Part 30.
<p>Council of Chief State School Officers (Washington, DC)</p>	<ul style="list-style-type: none"> • 2010- Focused on technical issues connected to implementing state-level accountability systems. Applied quantitative modeling to statewide assessment results to examine how potential business rule modifications might influence the distribution of accountability results. • 2007-Quality Assurance Process Management- Provided ASR members technical and practical solutions as a former state representative (Louisiana). Reviewed and contributed to ASR publications that focused on developing quality assurance procedures. • 2006- Taylor, R., Beaudoin, J.P., Auty, B., & Goldschmidt, P. (2006) <i>Validity Threats: Detection and control practices for state and local education officials</i>. Council for Chief State School Officers: Washington, D.C.
<p>Delaware Department of Education (Dover, DE)</p>	<ul style="list-style-type: none"> • 2012-Provided technical support to the DDOE regarding evidence needed for Peer Review. Created policy recommendations for the Secretary and Assistant Secretary of Education. Developed business rule changes (recommendations) necessary to integrate DCAS data into the state's transitional accountability system. • 2010-Designed the accountability system for Title III with the 2010 ACCESS for ELLs[®] results.

Beaudoin, John P.
Vitae 2015

	<ul style="list-style-type: none"> • 2012-Extracted, evaluated, and established Delaware’s <i>USED Standards and Assessment Peer Review</i> body of evidence in both 2005 and 2012. • 2006-Evaluated the use of accommodations for SWD and ELL students and established screening thresholds using multi-wave data.
Louisiana Department of Education	<ul style="list-style-type: none"> • 2002- Developed managed, and implemented Louisiana’s school and district accountability system in conjunction with implementing the new <i>No Child Left Behind Act of 2001</i> with LDE staff (53 staff members with a budget of approximately \$6 million). Provided technical and operational support for all Title I, Part A subgrantees (\$365 million) in the state, while conducting feasibility research for the LDE’s reading initiative (Reading First). Prepared and presented policy recommendations to the Louisiana’s Board of Elementary and Secondary Education (BESE). Conducted impact research (including Highly Qualified Teachers) regarding the statutory language in NCLB and represented the State Superintendent at national and legislative hearings. Created and promulgated the State’s school improvement efforts in conjunction with textbook adoptions and other Title I initiatives.
Maine Department of Education (Augusta, ME)	<ul style="list-style-type: none"> • 2008-Extracted, evaluated, and established Maine’s <i>USED Standards and Assessment Peer Review</i> body of evidence. Coordinated with the MDOE staff in producing and revising evidence associated with Maine’s High School Assessment (MHSA), which included the College Board’s SAT[®] and Measured Progress’s augmentation of the Math-A. • 2007-Developed the state’s Title III accountability design using empirical modeling. Developed and implemented the production cycle necessary to score and report Title III determinations for all school districts and consortia within the state using student-level data.
Office of the State Superintendent of Education (Washington, DC)	<ul style="list-style-type: none"> • 2008- Prepared the agency's evidence in science for submission to <i>USED Standards and Assessment Peer Review</i>. Completed a "white paper" for the agency in the comparability of English-language learner (ELL) accommodations on the 2008 DC CAS. Evaluated the use of accommodations for students with disabilities (SWD) on the 2008 DC CAS.
Nebraska Department of Education (Lincoln, NE)	<ul style="list-style-type: none"> • 2005- Developed the state’s Title III accountability system using empirical modeling. Developed and implemented the production cycle necessary to score and report Title III determinations for all school districts and consortia within the state. Produced Title III accountability results for the NDE from 2005 to 2009. Developed and published the agency’s Title III <i>Technical Manual for Title III Accountability</i>.
Oklahoma Department of Education (Oklahoma City, OK)	<ul style="list-style-type: none"> • 2006-Developed key components of ODOE NCES longitudinal data system grant application for use in measuring student academic growth over time.
Wyoming Department of Education (Cheyenne, WY)	<ul style="list-style-type: none"> • 2007-Provided technical assistance to the Wyoming Department of Education (WDE) regarding the agency’s alternate assessment. Conducted on-site briefings with WDE staff and co-developed an implementation plan used to address Section 4 of <i>USED Standards and Assessment Peer Review Guidance</i>.
Montana Office of Public Instruction (Helena, MT)	<ul style="list-style-type: none"> • 2006- Examined and provided consultation regarding the state's compliance review. Developed a series of recommendations to state education officials regarding <i>USED Standards and Assessment Peer Review</i>.

St. Tammany Parish School Board (Covington, LA)	<ul style="list-style-type: none"> • 2006-Developed on-line, curriculum-based assessments in grades 2-8 for use in measuring Louisiana's content. • 2004-Developed the assessment blueprint and specification tables necessary to coordinate item-development by the district. Implemented the Company's quality assurance reviews to ensure content alignment, developmental appropriateness, range of knowledge, and depth of knowledge prior to operational form construction. Created assessments for grades 2-7 in mathematics and reading.
St. John Parish School Board (Reserve, LA)	<ul style="list-style-type: none"> • 2004-Implemented the Company's quality assurance reviews to ensure content alignment, developmental appropriateness, range of knowledge, and depth of knowledge prior to operational form construction. Created operational assessments for grades 3-8 and conducted scoring activities.

Training

Pennsylvania Department of Education (Harrisburg, PA)	<ul style="list-style-type: none"> • 2013- General contractor for all student learning objectives (SLOs) development implementation tasks prior to the statewide launch. • 2012- Created the statewide SLO training series using the Homeroom learning platform. Integrated the SLO process into the greater educator effectiveness system. Created quality criteria for those performance measures used to measure student achievement within the SLO framework. • Trained over 650 PA educators, principals, Intermediate Units (IU), curriculum coordinators, PDE staff.
New Mexico Public Education Department (Santa Fe, NM)	<ul style="list-style-type: none"> • 2012-Trained over 70 NM educators using the <i>Assessment Literacy Series</i> process, modified this to produce six end-of-course assessments for use as alternate measures of competency for high school graduation in New Mexico. • 2013-Trained over 100 NM educators using the <i>Assessment Literacy Series</i> process, including the integration of performance measures. Developed 28 end-of-course assessments in the state's educator effectiveness system. Quality controlled all assessments using RIA's assessment rigor screening tool and quality assurance rubric, including evaluating the alignment of items/tasks to the New Mexico state content standards and Common Core State Standards. Provided technical recommendations to the NMPED senior staff on leveraging resources for greater access by NM educators.
Micosukee Tribe (Miami, FL)	<ul style="list-style-type: none"> • Spring 2013- Provided technical assistance in the development of the MIS alternate AYP definition as authorized within 25 C.F.R. Created business rules and integrated multiple assessment data into an overall school-based index score. Trained school improvement team and administrators using RIA's <i>Assessment Literacy Series</i>.
Maine Department of Education (Augusta, ME)	<ul style="list-style-type: none"> • 2011- Operated Maine's Summer Harvest Program/School and provided educational services to several hundred migrant students during the summers of 2011 and 2012. Developed a data validation process for multiple years necessary to support federal (CSPR) reports on migrant enrollment.
St. John Parish School Board (Reserve, LA)	<ul style="list-style-type: none"> • 2004- Implemented a seven-part professional development series to build teacher capacity in the construction and evaluation of classroom assessments. Developed scoring rubrics at the grade and classroom level to examine critical assessment characteristics for formative assessments used in grades K-8. Produced the Content Helper[®] for each teacher in the school. Distributed alignment tools within the materials which assisted teachers in developing formative assessments aligned to Louisiana's Grade Level Expectations.
St. James Parish School Board (Lutcher, LA)	<ul style="list-style-type: none"> • 2014 – Trained over 60 St. James Parish teachers using the Assessment Literacy Series program (Quick Start); aligned specification tables and blueprints to locally established scope and sequences (middle and high school). Developed 23 end-of-course assessments used by the district for pupil progression.

Evaluation

<p>Delaware Department of Education (Dover, DE)</p>	<ul style="list-style-type: none"> • 2013- Quality controlled all locally-developed assessments (for use in the state’s educator effectiveness system) using RIA’s assessment rigor screening tool and quality assurance rubric, including evaluating the alignment of items/tasks to the Delaware state content standards and Common Core State Standards. Provided technical support to local districts meeting the quality expectations adopted by the Delaware Department of Education. • Developed and implemented the Delaware Department of Education’s Internal Measures Project, which created high-quality, summative assessments for use in Delaware’s Educator Effectiveness System. <ul style="list-style-type: none"> ○ Trained over 700 DE educators using customized training modules that created and evaluated locally-developed assessments. ○ Developed the criteria procedures for evaluating vendor-made, student achievement measures (SAM-E). ○ Developed and implemented the training and quality standards for growth goals (student learning objectives-SLOs) of 47 content areas, including non-subject educators and professional staff (e.g., nurses). • Summer 2012- Reviewed over 550 assessments created using the assessment literacy process, including assessments in early childhood, 28 foreign languages, music, visual and performing art, physical education, health, and 95% of all courses offered in career and technical education in Delaware schools. • Summer 2013- Refined over 550 assessments created and field-tested in 2012, including assessments in early childhood, 28 foreign languages, music, visual and performing art, physical education, health, and 95% of all courses offered in career and technical education in Delaware schools. • Developed the first accountability technical manual (2004) in the nation that outlined the operational details associated with Delaware’s accountability system. Examined data quality practices associated with the Delaware Student Testing Program (DSTP), Delaware Student Information System (DELSIS), and targeted federal programs administered by the state. • Evaluated the state’s Accountability Growth Model design during its first year of implementation (Summer 2007). Conducted internal auditing of the agency’s processes used to make AYP determinations with the aforementioned Accountability Growth Model. • Extracted, evaluated, and established Delaware’s <i>USED Standards and Assessment Peer Review</i> body of evidence in both 2005 and 2012. • Evaluated the use of accommodations for SWD and ELL students (2006) and established screening thresholds using multi-wave data.
<p>New England Secondary School Consortium (Portland, ME)</p>	<ul style="list-style-type: none"> • 2014 – Published Annual Data Profile: 2012-13 and the NESSC’s Metrics: Procedural Guidebook (Fall 2014) • 2013- Created and published the NESSC’s <i>Annual Data Profile: SY 2011-12</i> (September 2013). Created and published the NESSC’s <i>Metrics: Procedural Guidebook</i> (February 2013).
<p>United States Department of the Interior - Bureau of Indian Education (Washington, D.C.)</p>	<ul style="list-style-type: none"> • Spring 2009- Trained BIE educators in using assessment data to determine their school’s AYP status. Created computational and resource support (in a web-based structure) necessary to calculate AYP for each of the 173 BIE-funded schools using 23 different, state-level, accountability models. Created technical guides for BIE-funded schools based upon their state’s accountability system.

<p>United States Department of Education (Washington, D.C.)</p>	<ul style="list-style-type: none"> • October 2012- Selected as a member of the ESEA Flexibility Peer Review team for “Window 3” of the ESEA Flexibility requests. Three requests for flexibility were reviewed. December 2011- Selected as a member of the ESEA Flexibility Peer Review team for “Window 1” of the ESEA Flexibility requests. • December 2008- Selected as the 2008 Chairperson for the USED Growth Model Peer Review process. Facilitated national experts’ review of several state growth models for use in making AYP determinations. States reviewed included: Texas, Colorado, Minnesota, District of Columbia, Pennsylvania, North Dakota, and New York. April 2008- Participated in the technical evaluation of state growth model proposals for use in making AYP determinations in accordance with USED guidelines. States reviewed included: Missouri, New Mexico, Pennsylvania, District of Columbia, Michigan, and Minnesota. • January 2005 to 2012- Selected in the initial cohort of Peer Reviewers (2005) to review evidence associated with state content standards and large-scale assessments. Evaluated evidence presented by state agencies against the critical elements found within the <i>USED’s Peer Review Guidance</i> document. States reviewed included: Indiana, Illinois, Iowa, New Mexico, Puerto Rico, South Carolina, Tennessee, Texas, North Carolina, Hawaii, Wyoming, New York, New Jersey, Alaska, Nevada, and the Mariana Islands.
<p>Maine Department of Education Augusta, ME)</p>	<ul style="list-style-type: none"> • Summer 2008- Conducted IT audits of MEDMS data used to calculate the on-time graduation rate (i.e., as required by 34 C.F.R. 200.19). Developed QA/QC risk management document and implemented procedures in partnership with the MDOE staff.
<p>Nebraska Department of Education (Lincoln, NE)</p>	<ul style="list-style-type: none"> • Spring 2008- Evaluated accountability business rules with end-user guidelines during initial piloting of the state’s new student information system. Supported NDE officials in developing a multi-million dollar NCES grant to implement a longitudinal data system. Developed a conceptual framework and the RFP for the NSSRS Decision Support System (DSS).
<p>St. Tammany Parish School Board (Covington, LA)</p>	<ul style="list-style-type: none"> • Spring 2003- Extracted student performance data from the statewide assessment across multiple years to produce school and district-level Data Notebooks (SY 2004 thru 2008). Data Notebooks address multi-level performance goals by providing quantitative trend data to non-technical audiences. Aligned the business rules within the Strategic Plan, School Improvement Plan, and Data Notebooks to ensure comparability.
<p>McComb School District (McComb, MS)</p>	<ul style="list-style-type: none"> • Conducted a comprehensive needs assessment for each school within the district. Organized data into school and district-level reports to understand how the district’s reform initiatives were being actualized.

Christy Lyn Hovanetz, Ph.D.

7326 Willow Lane, Brooklyn Park, MN 55444 • 850-212-0243 • ChristyHovanetz@gmail.com

Professional Profile:

- Experience developing public education policy and programs
- Leadership experience in the public and private sector managing education projects and teams
- Presentation experience to local, state, national, and international audiences, State Boards of Education, Legislatures, Governors, superintendents, administrators, teachers, students, and public
- Teaching licensure and experience in the K-12 public schools
- Skilled at policy analysis, data analysis, research and evaluation, accountability and assessment, program development and implementation, technical writing, communication with stakeholders

Education:

Askew School of Public Administration and Policy - Florida State University, Tallahassee, FL.

Doctor of Philosophy. Pi Alpha Alpha Honor Society. Graduated in December 2005.

Concentration: Public Policy, specializing in education policy.

Dissertation Title: Education Policy Adoption at the Local Governance Level: An Empirical Analysis of Factors Influencing the Adoption of Reading First in Florida School Districts.

<http://etd.lib.fsu.edu/theses/available/etd-10272005-211856/>

Major Professor: Frances Stokes Berry.

Hubert H. Humphrey Institute of Public Affairs - University of Minnesota, Minneapolis, MN.

Master of Public Affairs. Graduated in December 1999.

Concentrations: Public and Nonprofit Leadership and Management and Conflict Management.

Adviser: John E. Brandl

St. Cloud State University, St. Cloud, MN.

Bachelor of Science. Kappa Delta Pi Honor Society. Graduated summa cum laude in May 1998.

Major/Minor: Education and Mathematics.

University Honors Program Graduate.

Professional Experience:

Foundation for Excellence in Education, Tallahassee, FL.

Senior Policy Fellow, November 2008 to Present.

Lead the national education policy research, data, and evaluation agenda for accountability at the Foundation, a 501(c)3 not-for-profit education policy and research organization founded by Jeb Bush, former Governor of Florida. Support states implementing evidence-based education reform by developing a customized plan of action and supporting research that include policy papers, data analysis, presentations and testimony, and provide technical support to ensure successful implementation.

Summit Education Associates LLC, Minneapolis, MN.

Owner/President, July 2010 to Present.

President and founder of an education consultancy which provides policy and data services to clients. The company focuses on providing clients with specialized services in education policy, data analysis, evaluation and research, strategy, project management and program implementation.

Minnesota Department of Education, State of Minnesota, Roseville, MN.

Assistant Commissioner, March 2007 to November 2008.

Led the Research and Assessment Office responsible for K-12 statewide assessment program; No Child Left Behind, including school accountability, federal title programs and policy; and the School Improvement Office, including Quality Compensation for effective teacher performance, statewide professional development, and school improvement planning. Responsible for leading the research and evaluation efforts for the Department.

Florida Department of Education, State of Florida, Tallahassee, FL.

Assistant Deputy Commissioner, Accountability, Research, and Measurement Division, July 2006 to March 2007.

Led the Evaluation and Reporting Office (described below) and the Education Information and Accountability Services Office which is responsible for K-12 education data collection. Data is used for funding, reporting, evaluation, and policy making. Collecting data, improving data reporting accuracy by school districts, responding to data inquiries, fulfilling reporting requirements, and reviewing and developing data collection procedures are primary responsibilities.

Director, Evaluation and Reporting Office, November 2003 to July 2006.

Lead the office responsible for public school accountability and No Child Left Behind Act policy in Florida. Calculate school grades under Florida's A+ Plan and Adequate Yearly Progress under the federal No Child Left Behind Act of 2001. Developing accountability systems for Florida's voluntary pre-kindergarten program, supplemental educational services provide in Florida as required under No Child Left Behind, and Florida's teacher pay for performance initiative. Perform research, evaluation, and policy and data analysis related to PreK-12 education programs and policy initiatives.

Director of Reading First, Just Read, Florida! Office, February 2002 to November 2003.

Co-authored and directed Florida's Reading First grant under the No Child Left Behind Act of 2001. Secured funding and was responsible for \$300 million budget, policy development, and implementation of Reading First. Led team of academics and education practitioners in the implementation of Reading First in Florida. Presented and provided technical assistance to K-20 education professionals regarding the Governor's Just Read, Florida! initiative and Reading First.

Executive Office of the Governor, Office of Policy and Budget, State of Florida, Tallahassee, FL.

Governmental Analyst, July 2000 to February 2002.

Wrote legislation and amendments to current law for the state of Florida based on the Governor's priority education initiatives. Researched and analyzed current and prospective legislation for policy and fiscal impact. Served as a researcher for the Governor, appointed task forces, and education partnerships.

Office of Program Policy Analysis & Government Accountability, State of Florida, Tallahassee, FL.

Legislative Policy Analyst, K-12 Education Policy, September 1999 to July 2000.

Performed objective, independent research and analysis of Florida's K-12 public school system. Analysis included comparison of statutory performance measures, analysis of program productivity, consideration of program value, budget analysis, and made recommendations for future effectiveness. Conducted interviews, created and administered surveys, performed legislative research, quantitative and qualitative research, and produced detailed reports containing recommendations to high-level policymakers.

Academic Experience:

Florida State University, Tallahassee, FL.

Adjunct Professor. PAD 3003, Public Administration in Society.

Honors and Leadership:

- Peer Reviewer for Elementary and Secondary Education Act state waiver requests, United States Department of Education (2011, 2013, and 2015)
- Peer Reviewer Institutes for Education Sciences Grants (2014)
- American Marshall Memorial Fellow, German Marshall Fund of the United States (2011)
- Policy Fellow at the Hubert H. Humphrey Institute of Public Affairs (2009-10)
- American Delegate to Nicaragua and El Salvador, American Council of Young Political Leaders, U.S. Department of State (2008)
- Mentor, Hubert H. Humphrey Institute of Public Affairs, University of Minnesota
- Pi Alpha Alpha, the National Honor Society for Public Affairs and Administration
- Kappa Delta Pi, International Honor Society in Education
- Junior League of Tallahassee
- Commissioner's Leadership Seminar, Florida Department of Education
- YMCA youth soccer coach
- St. Cloud State University, Student Government President (1997-98)
- St. Cloud State University, NCAA Division II Women's Soccer, Player
- Delta Zeta International Sorority, President (1996-97)