

Stefanie A. Wind
The University of Alabama
College of Education
Educational Measurement and Evaluation
Department of Educational Studies in Psychology, Research Methodology, and Counseling
swind@ua.edu
<https://professorwind.science>

EDUCATION

- Emory University**, PhD in Educational Measurement 2014
Dr. George Engelhard, Jr., Advisor
Dr. Yuk Fai Cheong and Dr. Robert J. Jensen, Dissertation committee members
Dissertation Title: *Evaluating Rater-Mediated Assessments with Rasch Measurement Theory and Mokken Scaling*
- Emory University**, MA in Educational Studies: Quantitative Methodology 2012
Measurement, Assessment, and Evaluation: Major field of study
Dr. George Engelhard, Jr., Advisor
Dr. Yuk Fai Cheong and Dr. Robert J. Jensen, Thesis committee members
Thesis title: *Evaluating the Quality of Ratings in Writing Assessment: Rater Agreement, Error, and Accuracy*
- University of West Florida**, BA in Elementary Education 2009
Elementary Education (K – 6 Certification), ESOL and Reading specializations, *Summa cum Laude*
- University of West Florida**, BM in Music Performance 2009
Harp Performance, *Summa cum Laude*

ACADEMIC APPOINTMENTS

- The University of Alabama** August 2015 – present
Assistant Professor of Educational Measurement and Evaluation
College of Education
Department of Educational Studies in Psychology, Research Methodology, and Counseling
- Georgia Institute of Technology** June 2014 – June 2015
Research Scientist II
College of Sciences
Center for Education Integrating Science, Mathematics, and Computing (CEISMC)
- Emory University** September 2014 – May 2015
Postdoctoral Research Fellow
College of Arts and Sciences, Division of Educational Studies

LINES OF RESEARCH

1. Exploration of methodological issues in the field of psychometrics, with emphases on: (a) rating scales; (b) polytomous measurement models; (c) rating quality indicators; (d) Mokken scaling; and (e) Rasch measurement theory.
2. Application of psychometric methods (especially those included in Line 1) to substantive areas related to education.

PEER-REVIEWED JOURNAL PUBLICATIONS

(* Indicates student author)

Wind, S. A. (2018). Digital ITEMS module 03: Nonparametric item response theory. *Educational Measurement: Issues and Practice*, 37(4), 63-64. <http://dx.doi.org/10.1111/emip.12239>

Wind, S. A., Alemdar, M., Lingle, J. A., Moore, R., & Asilkalkan, A.*.(in press). Exploring student understanding of the engineering design process using distractor analysis. *International Journal of STEM Education*.

Wind, S. A., & Walker, A. A. (in press). Exploring the correspondence between traditional score resolution methods and person fit indices in rater-mediated writing assessments. *Assessing Writing*

Wind, S. A., & Jones, E. (in press). The effects of incomplete rating designs in combination with rater effects. *Journal of Educational Measurement*.

Wind, S. A. & Sebok-Syer, S. S. (in press). Examining differential rater functioning using a between fit approach. *Journal of Educational Measurement*.

Pennington, C. G.* , Curtner-Smith, M. D., & **Wind, S. A.** (in press). Impact of a physical education teacher's age on elementary student's perceptions of effectiveness and learning. *Journal of Teaching in Physical Education*.

Pennington, C. G.* , Curtner-Smith, M. D., & **Wind, S. A.** (in press). Impact of a physical education teacher's age on high school pupil's perceptions of effectiveness and learning. *European Physical Education Review*.

Peabody, M. & **Wind, S. A.** (in press). Exploring the influence of judge proficiency on standard-setting judgments for medical examinations. *Journal of Educational Measurement*.

- Holmes, C. J.*, **Wind, S. A.** (2018). Heart rate variability responses to an undulating resistance training program in free-living conditions: A case study in a collegiate athlete. *Sports*, 6(121), 1-12. <https://doi.org/10.3390/sports6040121>
- Peabody, M., & **Wind, S. A.** (in press). Exploring the stability of differential item functioning across administrations and critical values using the Rasch separate calibration t-test method. *Measurement: Interdisciplinary Research and Perspectives*.
- Gale, J. D., Koval, J., Ryan, M., Usselman, M., & **Wind, S. A.** (in press). Implementing NGSS engineering disciplinary core ideas in middle school science classrooms: Results from the field. *Journal of Pre-College Engineering Education Research*.
- Wind, S. A.** (2018). Using Guttman errors to explore rater fit in rater-mediated performance assessments. *Methodological Innovations*, 11(3), 1-17. <https://doi.org/10.1177/2059799118814396>
- Wind, S. A.** & Stager, C.* (in press). The impacts of characteristics of disconnected subsets on group anchoring in incomplete rater-mediated assessment networks. *Psychological Test and Assessment Modeling*.
- Wind, S. A.**, Mansouri, B.*, & Jami, P. Y.* (in press). Examining the psychometric properties of a questionnaire to investigate Iranian EFL learners' preferences for isolated and integrated form-focused grammar instruction. *Journal of Applied Measurement*.
- Wind, S. A.** (2018). Examining the impacts of rater effects in performance assessments. *Applied Psychological Measurement*. Advance online publication. <https://doi.org/10.1177/0146621618789391>
- Hua, C.*, & **Wind, S. A.** (2018). Exploring the psychometric properties of the Mind Map Scoring Rubric. *Behaviormetrika*. Advance online publication. <https://doi.org/10.1007/s41237-018-0062-z>
- Wind, S. A.**, Jami, P. Y.* & Mansouri, B.* (2018). Exploring the psychometric properties of the Empathy Quotient for Farsi speakers. *Current Psychology*. Advance online publication. <https://doi.org/10.1007/s12144-018-9938-z>
- Wind, S. A.** & Schumacker, R. E. (2018). Exploring within-rater category ordering: A simulation study using adjacent-categories Mokken scale analysis. *Educational and Psychological Measurement*, 78(5), 887-904. <https://doi.org/10.1177/0013164417724841>
- Jones, E., & **Wind, S. A.** (2018). Using repeated rating occasions to improve measurement precision in incomplete rating designs. *Journal of Applied Measurement*, 19(2), 148-161.
- Alemdar, M., Lingle, J. A., **Wind, S. A.**, & Moore, R. (2017). Developing an engineering design assessment using think-aloud interviews. *International Journal of Engineering Education*, 33(1B), 441-452.

- Wind, S. A., Tsai, C-L., Grajeda, S. B., & Bergin, C. (2018).** Principals' use of rating scale categories in teacher evaluation. *School Effectiveness and School Improvement*, 29(3), 485-510.
<https://doi.org/10.1080/09243453.2018.1470989>
- Cook, R., McKibben, B., & **Wind, S. A. (2018).** Supervisee perception of power in clinical supervision: The Power Dynamics in Supervision Scale. *Training and Education in Professional Psychology*, 12(3), 188-195. <https://doi.org/10.1037/tep0000201>
- Wind, S. A. & Jones, E. (2018).** Exploring the influence of range restrictions on connectivity in sparse assessment networks: An illustration and exploration within the context of classroom observations. *Journal of Educational Measurement*, 55(2), 217-242.
<https://doi.org/10.1111/jedm.12173>
- Wind, S. A., Ooi, P. S., & Engelhard, G. (2018).** Exploring decision consistency across rating designs in rater-mediated music performance assessments. *Musicae Scientiae*. Advance online publication.
<https://doi.org/10.1177/1029864918761184>
- Wind, S. A., & Jones, E. (2018).** The stabilizing influences of linking set size and model-data fit in incomplete rater-mediated assessment networks. *Educational and Psychological Measurement*, 78(4), 679-707. <https://doi.org/10.1177/0013164417703733>
- Wind, S. A., & Patil, Y. J.* (2018).** Exploring incomplete rating designs with Mokken scale analysis. *Educational and Psychological Measurement*, 78(2), 319-342.
<https://doi.org/10.1177/0013164416675393>
- Wang, J., Engelhard, G., Jr., & **Wind, S. A. (2018).** A tale of two models: Psychometric and cognitive perspectives on rater-mediated assessments using accuracy ratings. *Psychological Test and Assessment Modeling*, 60(1), 33-52.
- Wind, S. A. & Wesolowski, B. C. (2018).** Evaluating differential rater accuracy over time in solo music performance assessment. *Bulletin of the Council of Research in Music Education*, 215, 33-55.
<https://doi.org/10.5406/bulcouresmusedu.215.0033>
- Knisely, K. A., & **Wind, S. A. (2017).** Developing a survey to explore sense of belongingness related to language learning using Rasch measurement theory. *Pensamiento Educativo: Journal of Latin-American Educational Research, (Special Issue on Applied Educational Measurement)*, 54(2), 1-20.
- Wind, S. A., & Engelhard, G., Jr. (2017).** Exploring rater errors and systematic biases using adjacent-categories Mokken models. *Psychological Test and Assessment Modeling*, 59(4), 493-515.
- Wind, S. A., Stager, C*., & Patil, Y. J.* (2017).** Exploring the relationship between essay features and rating quality for L1 and L2 writers. *Assessing Writing*, 34, 1-15.
<https://doi.org/10.1016/j.asw.2017.08.003>

- Wind, S. A.,** Wolfe, E. W., Engelhard, G., Jr., Foltz, P., & Rosenstein, M. (2017). The influence of rater effects in training sets on the psychometric quality of automated scoring for writing assessments. *International Journal of Testing*, 18(1), 27-49. <https://doi.org/10.1080/15305058.2017.1361426>
- Wesolowski, B. C., & **Wind, S. A.** (2017). Investigating rater accuracy in the context of secondary-level solo instrumental music performance. *Musicae Scientiae*. Advance online publication. <https://doi.org/10.1177/1029864917713805>
- Wind, S. A.,** & Schumacker, R. E. (2017). Detecting measurement disturbances in rater-mediated assessments. *Educational Measurement: Issues and Practice*, 36(4), 44-51. <https://doi.org/10.1111/emip.12164>
- Bergin, C., **Wind, S. A.,** Grajeda, S., & Tsai, C.-L. (2017). Teacher evaluation: Are principals' classroom observations accurate at the conclusion of training? *Studies in Educational Evaluation*, 55, 19-26. <https://doi.org/10.1016/j.stueduc.2017.05.002>
- Wind, S. A.** (2017). An instructional module on Mokken Scale Analysis. *Educational Measurement: Issues and Practice*, 36(2), 50-66. <https://doi.org/10.1111/emip.12153>
- Patterson, B. F., **Wind, S. A.,** & Engelhard, G., Jr. (2017). Incorporating expert ratings into model-based rater monitoring procedures using signal detection theory. *Applied Psychological Measurement*, 41(6), 472-491. <https://doi.org/10.1177/0146621617698452>
- Wind, S. A.** & Peterson, M. E.* (2017). A systematic review of methods for evaluating rating quality in language assessment. *Language Testing*, 35(2), 161-192. <https://doi.org/10.1177/0265532216686999>
- Wind, S. A.,** Alemdar, M., Gale, J. D., Lingle, J. A., & Moore, R. (2017). Developing an engineering design process assessment using mixed methods: An illustration with Rasch measurement theory and cognitive interviews. *Journal of Applied Measurement*, 18(2), 100-121.
- Wesolowski, B. W., **Wind, S. A.,** & Engelhard, G. Jr. (2017). Evaluating differential rater functioning over time in the context of solo music performance assessment. *Bulletin of the Council of Research in Music Education*, 212, 75-98. <https://doi.org/10.5406/bulcouresmusedu.212.0075>
- Gale, J. D., Koval, J., **Wind, S. A.,** Ryan, M., & Usselman, M. (2017). Exploring student understanding of force and motion using a simulation-based performance assessment. *Journal of Research in STEM Education*, 12(1), 39-58.
- Wind, S. A.** (2017). Adjacent-categories Mokken models for rater-mediated assessments. *Educational and Psychological Measurement*, 77(2), 330-350. <https://doi.org/10.1177/0013164416643826>
- Gale, J. D., **Wind, S. A.,** Koval, J., Dagosta, J., Ryan, M., & Usselman, M. (2016). Simulation-based performance assessment: An innovative approach to exploring understanding of physical science

concepts. *International Journal of Science Education*, 38(14), 2284-2302.
<https://doi.org/10.1080/09500693.2016.1236298>

- Wind, S. A.** (2016). Examining the psychometric quality of multiple-choice assessment items using Mokken scale analysis. *Journal of Applied Measurement*, 17(2), 142–165.
- Wind, S. A.**, Engelhard, G., & Wesolowski, B. (2016). Exploring the effects of rater linking designs and rater fit on achievement estimates within the context of music performance assessments. *Educational Assessment*, 21(4), 278–299. <https://doi.org/10.1080/10627197.2016.1236676>
- Wesolowski, B., **Wind, S. A.**, & Engelhard, G., Jr. (2016). Examining rater precision in music performance assessment: An analysis of rating scale structure using the multifaceted Rasch partial credit model. *Music Perception*, 33(5), 662-678.
<https://doi.org/10.1525/mp.2016.33.5.662>
- Wind, S. A.**, & Engelhard, G., Jr. (2015). Exploring rating quality in rater-mediated assessments using Mokken scaling. *Educational and Psychological Measurement*, 76(4), 685-706. <https://doi.org/10.1177/0013164415604704>
- Wind, S. A.** (2015). Evaluating the quality of analytic ratings with Mokken scaling. *Psychological Test and Assessment Modeling*, 57(3), 423-444.
- Wind, S. A.**, & Gale, J. D. (2015). Diagnostic opportunities with distractor-driven multiple-choice items in the context of a physical science assessment. *Science Education*, 99, 721-741. <https://doi.org/10.1002/sce.21172>
- Wesolowski, B., **Wind, S. A.**, & Engelhard, G. Jr. (2015). Rater fairness in music performance assessment: Evaluating model-data fit and differential rater functioning. *Musicae Scientiae*, 19(2), 147-170. <https://doi.org/10.1177/1029864915589014>
- Knisely, K. A., & **Wind, S. A.** (2015). Exploring gendered language attitudes using Rasch measurement theory. *Journal of Applied Measurement*, 16(1), 95-112.
- 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. <https://doi.org/10.1080/15305058.2014.931281>
- Wind, S. A.** (2014). Guidelines for rating scales based on Rasch measurement theory and Mokken scaling. *Journal of Applied Measurement*, 15(2), 100-133.
- Wind, S. A.** & Engelhard G., Jr. (2013). How invariant and accurate are domain ratings in writing assessment? *Assessing Writing*, 18(4), 278-299. <https://doi.org/10.1016/j.asw.2013.09.002>
- Engelhard, G., Jr., & **Wind, S. A.** (2013). Educational testing and schooling: Unanticipated consequences of purposive social action. *Measurement: Interdisciplinary Research and Perspectives*, 11, 30-35. <https://doi.org/10.1080/15366367.2013.784156>

Kaliski, P., **Wind, S. A.**, Engelhard, G., Jr., Morgan, D., Reshetar, R., & Plake, B. (2013). Using the Many-Faceted Rasch Model to evaluate standard-setting judgments: Setting performance standards for Advanced Placement examinations. *Educational and Psychological Measurement*, 73(2), 1-26. <https://doi.org/10.1177/0013164412468448>

Wind, S. A. & Engelhard, G., Jr. (2012). Evaluating the quality of ratings in writing assessment: Rater agreement, error, and accuracy. *Journal of Applied Measurement*. 13(4), 1-15.

BOOKS

Engelhard, G., Jr., & **Wind, S. A.** (2018). *Invariant measurement with raters and rating scales: Rasch models for rater-mediated assessments*. New York, NY: Taylor and Francis.

BOOK CHAPTERS

Wind, S. A. (2018). Lexiles. In B. Fey (Ed.) *The SAGE Encyclopedia of Educational Research, Measurement and Evaluation*, p. 973. Los Angeles, CA: Sage.

Wesolowski, B. W. & **Wind, S. A.** (in press). Validity, reliability, and fairness. In T. Brophy (Ed.) *The Oxford Handbook of Assessment Policy & Practice in Music Education*.

PEER-REVIEWED RESEARCH NOTES

Wind, S. A., Hua, C.*, Porter, M.*, Stager, C.*, & Zhang, S.* (2017). An updated version of Ben Wright's method and meaning of measurement infographic. *Rasch Measurement Transactions*, 30(3), 1595-1597.

Wind, S. A., & Engelhard, G., Jr. (2014). Model-data fit and adjustments for rater effects. *Rasch Measurement Transactions*, 28(3), 1483-1484.

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

Wind, S. A., & Engelhard, G., Jr. (2011). Rater-mediated domain response functions. *Rasch Measurement Transactions*, 25(2), 1321-1322.

PEER-REVIEWED TECHNICAL REPORTS

Wind, S. A. (2017). *Examining student achievement outcomes based on multiple-choice assessments for AMP-IT-UP*. (CEISMC-TR-1706a). Atlanta, GA: Georgia Institute of Technology.

Wind, S. A. (2016). *Examining student achievement outcomes related to Engineering Design: Results from Fall 2015*. (CEISMC-TR-1607a). Atlanta, GA: Georgia Institute of Technology.

- Wind, S. A.** (2016). *Examining student achievement outcomes related to Engineering Design: Results from Spring 2016.* (CEISMC-TR-1607b). Atlanta, GA: Georgia Institute of Technology.
- Wind, S. A.** (2016). *Examining student achievement outcomes related to AMP-IT-UP Supplementary Modules: Results from the academic year 2015-2016.* (CEISMC-TR-1607c). Atlanta, GA: Georgia Institute of Technology.
- Wind, S. A.** (2016). *Examining student achievement outcomes related to AMP-IT-UP Supplementary Modules: Results from Spring 2016.* (CEISMC-TR-1607d). Atlanta, GA: Georgia Institute of Technology.
- Wind, S. A.** (2014). *Examining student achievement outcomes in disciplinary science content, science practices, and engineering concepts using pre- and post-assessments for the slider curriculum: Results from the academic year 2013-2014.* (CEISMC-TR-1406). Atlanta, GA: Georgia Institute of Technology.
- Wind, S. A.** (2014). *Exploring item subsets in SLIDER multiple-choice assessments: Results from the academic year 2013-2014 (CEISMC-TR-1408).* Atlanta, GA: Georgia Institute of Technology.
- Song, T., **Wind, S. A.**, & Wolfe, E. W. (2013). *Rater performance index adverse impact study for the performance scoring center.* (Pearson Technical Report). Iowa City, IA: Pearson.
- Engelhard, G., Jr. & **Wind, S. A.** (2013). *Rating quality studies using Rasch Measurement Theory.* (College Board Research and Development Report No. 2013-3). New York, NY: College Board.
- Engelhard, G., Jr., **Wind, S. A.**, Kobrin, J., & Chajewski, M. (2013). *Differential item and person functioning in large-scale writing assessments within the context of the SAT reasoning test.* (College Board Research and Development Report No. 2013-6). New York, NY: College Board.

MANUSCRIPTS UNDER REVIEW IN PEER-REVIEWED JOURNALS

(* Indicates student author)

- Wind, S. A.** (Under second review: Revise and resubmit). A nonparametric procedure for exploring differences in rating quality across student subgroups based on adjacent-categories Mokken Models.
- Hodges, T., Wright, K. L., **Wind, S. A.**, Matthews, S., Zimmer, W. K.*, McTigue, E. M. Developing and validating the writing rubric to inform teacher educators (WRITE). (Under second review: Revise and resubmit).
- Wesolowski, B., & **Wind, S. A.** (Under review). A three-model framework to quantitatively investigate rater variability in the measurement of latent musical constructs.
- Walker, A. A. & **Wind, S. A.** (Under review). Identifying misfitting achievement estimates in performance assessments: An illustration using Rasch and Mokken scale analyses.

- Zhang, S.* & **Wind, S. A.** (Under review). Cross-cultural comparisons of school leadership using Rasch measurement.
- Wind, S. A.** & Wesolowski, B. C. (Under review). Are music performance classroom assessments psychometrically suitable to infer student achievement or teacher effectiveness? A nonparametric case study using Mokken scale analysis.
- Wind, S. A.** (Under review). Applying Mokken scaling techniques to incomplete rating designs for educational performance assessments
- Pennington, C. G.*, Curtner-Smith, M. D., & **Wind, S. A.** (Under review). Impact of a physical education teacher's age on middle school student's perceptions of effectiveness and learning.
- Lingle, J. A., Alemdar, M., **Wind, S. A.**, & Moore, R. (Under review). Examining the relationships between student knowledge of the engineering design process and 21st century skills in pre-college engineering courses.
- Wind, S. A.**, Jones, E., & Bergin, C. (Under review). The effects of teacher and principal characteristics on evaluation of teaching effectiveness.

PUBLISHED CONFERENCE PROCEEDINGS

- Wesolowski, B. C., **Wind, S. A.**, & Engelhard, Jr., G. (2016). Rater analyses in music performance assessment: Application of the Many Facet Rasch Model. In T.S. Brophy, J. Marlatt, & G. K. Ritcher (Eds.), *Connecting practice, measurement, and evaluation: Selected papers from the 5th International Symposium on Assessment in Music Education*, (pp. 335-356). Chicago, IL: GIA.

PAPERS PRESENTED AT PEER-REVIEWED CONFERENCES

(* Indicates student author)

- Wind, S. A.** (2018, July). Using Mokken scaling techniques to evaluate educational assessments. Paper presented at the European Association of Methodology Conference in Jena, Germany.
- Wind, S. A.** (2018, April). Considering the implications of cumulative and adjacent-categories models for raters: An illustration using Mokken Scale Analysis. Paper presented at the meeting of the National Council on Measurement in Education, New York, New York.
- Wind, S. A.**, Jones, E., Bergin, C. & Jensen, K. (2018, April). The effects of principal and teacher characteristics on evaluation of teaching effectiveness. Paper presented at the meeting of the American Educational Research Association, New York, New York.

- Jones, E., & **Wind, S. A.** (2018, April). Using repeated rating occasions to improve measurement precision in incomplete rating designs. Paper presented at the meeting of the National Council on Measurement in Education, New York, New York.
- Wind, S. A.** (2018, April). Examining the practical consequences of rater effects in performance assessments. Paper presented at the meeting of the American Educational Research Association, New York, New York.
- Zhang, S.*, **Wind, S. A.**, & Porter, M.*. (2018, April). Cross-cultural comparisons of school leadership using Rasch measurement. Paper presented at the meeting of the American Educational Research Association, New York, New York.
- Wind, S. A.**, Alemdar, M., Lingle, J. A., Newton, S., Gale, J., Moore, R., & Asilkalkan, A.* (2018, April). Exploring student understanding of the engineering design process using distractor analysis. Paper presented at the meeting of the American Educational Research Association, New York, New York.
- Wind, S. A.** & Walker, A. A. (2018, April). Exploring the correspondence between traditional score resolution methods and person fit indices in rater-mediated writing assessments. Paper presented at the meeting of the International Objective Measurement Workshop, New York, New York.
- Wind, S. A.** (2018, April). Using Guttman errors to explore rater fit in rater-mediated performance assessments. Paper presented at the meeting of the International Objective Measurement Workshop (Plenary Session), New York, New York.
- Hua, C.* & **Wind, S. A.** (2018, April). Exploring the psychometric properties of the Mind Map Scoring Rubric. Paper presented at the meeting of the International Objective Measurement Workshop, New York, New York.
- Cook, R., McKibben, W. B., & **Wind, S. A.** (2017, September). Examining power dynamics in supervision: The psychometric properties of the PDSS. Association for Assessment and Research in Counseling Conference, Pheonix, AZ.
- Hua, C.* Stager, C.*, & **Wind, S. A.** (2017, September). Synthesis of articles in the Journal of Applied Measurement: 2000-2016. Paper presented at the meeting of the International Outcomes Measurement Conference, Chicago, Illinois.
- Wind, S. A.** & Sebok-Syer, S. (2017, May). Infit, outfit, and between fit statistics for raters: The difference is in the details. Paper presented at the annual meeting of the Canadian Society for the Study of Education, Toronto, Ontario, Canada.
- Wind, S. A.**, Engelhard, G., Jr., & Ooi, P. S. (2017, April). Exploring decision consistency across rating designs in rater-mediated assessments. Paper presented at the annual meeting of the American Educational Research Association, San Antonio, Texas.

- Wind, S. A.**, Tsai, C-L., Chapman, S. B., & Bergin, C. (2017, April). Principals' use of rating scale categories in teacher evaluation. Paper presented at the annual meeting of the American Educational Research Association, San Antonio, Texas.
- Wind, S. A.**, & Stager, C. G.* (2017, April). Exploring the relationship between essay features and rating quality for L1 and L2 writers. Paper presented at the annual meeting of the American Educational Research Association, San Antonio, Texas.
- Walker, A. A. & **Wind, S. A.** (2017, April). Detecting and interpreting misfitting achievement estimates in performance assessments: an illustration using Rasch and Mokken scale analyses. Paper presented at the annual meeting of the American Educational Research Association, San Antonio, Texas.
- Wind, S. A.** & Schumacker, R. E. (2017, April). Exploring within-rater category ordering: A simulation study using adjacent-categories Mokken scale analysis. Paper presented at the annual meeting of the National Council on Measurement in Education, San Antonio, Texas.
- Wind, S. A.**, & Patil, Y. J.* (2016, July). Exploring incomplete rating designs with Mokken scale analysis. Paper presented at the annual meeting of the International Test Commission, Vancouver, Canada.
- Wind, S. A.**, Engelhard, G., Jr., & Raczynski, K. (2016, April). Exploring the relationship between essay features and rating quality. Paper presented at the annual meeting of the International Objective Measurement Workshop, Washington, DC.
- Wind, S. A.** & Engelhard, G., Jr. (2016, April). Exploring rater errors and systematic biases in language assessment using Mokken scale analysis. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.
- Peabody, M. & **Wind, S. A.** (2016, April). Exploring the influence of judge proficiency on standard-setting judgments for medical examinations. Paper presented at the annual meeting of the National Council on Measurement in Education, Washington, DC.
- Gale, J. G., Koval, J. & **Wind, S. A.** (2016, April). Exploring student understanding of force and motion using a simulation-based performance assessment. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.
- Bergin, C., **Wind, S. A.**, Chapman, S. B., & Tsai, C-L. (2016, April). Teacher evaluation: Are principals' classroom observations accurate? Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.
- Lingle, J. A., Alemdar, M., **Wind, S. A.**, Moore, R. & Usselman, M. (2016, April). The development and validation of an engineering assessment. Paper presented at the annual meeting of the National Association on Research for Science Teaching, Baltimore, MD.

- Peterson, M. E.*, & **Wind, S. A.** (2016, March). Rating the ratings: Methods for evaluating rating quality in language assessment. Poster presented at the Educational Research in Psychology, Research Methodology, and Counseling Symposium at the University of Alabama, Tuscaloosa, AL.
- Wind, S. A.** (2015, April). Examining the psychometric quality of multiple-choice assessment items with Mokken scale analysis. Paper presented at the meeting of the International Outcomes Measurement Conference, Chicago, IL.
- Wind, S. A.** (2015, April). Evaluating the quality of analytic ratings with Mokken scaling. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL.
- Wind, S. A.**, Wesolowski, B., & Engelhard, G., Jr. (2015, April). Exploring the effects of rater linking designs and rater fit on achievement estimates within the context of music performance assessments. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Wind, S. A.**, Alemdar, M., Gale, J. D., Lingle, J., & Moore, R. (2015, April). Developing an engineering design process assessment using mixed methods: An illustration with Rasch measurement theory and cognitive interviews. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Wesolowski, B., **Wind, S. A.**, & Engelhard, G. Jr. (2015, February). Assessing jazz big band performances with the Many-Facet Rasch Model. Paper presented at the International Symposium on Assessment in Music Education conference, Williamsburg, VA.
- Cheong, Y. F., **Wind, S. A.**, Engelhard, G., Jr., & Kaliski, P. (2014, August). A comparison of the Rasch and the hierarchical cross-classified modeling approach to standard setting ratings within the context of the AP World History Examination. Paper presented at the annual meeting of the Pacific Rim Objective Measurement Society, Guangzhou, People's Republic of China.
- Wind, S. A.**, Wolfe, E. W., Engelhard, G., Jr., Foltz, P., & Rosenstein, M. (2014, July). Influence of rater effects on the training of automated scoring engines. Paper presented at the annual meeting of the International Test Commission (ITC), San Sebastian, Spain.
- Wind, S. A.** (2014, April). Diagnosing rating scale effectiveness with Rasch models and Mokken scaling. Paper presented at the biannual meeting of the International Objective Measurement Workshop, Philadelphia, PA. **(Received Graduate Student Award)**.
- Wind, S. A.** & Gale, J. D. (2014, April). Diagnostic opportunities with distractor-driven multiple-choice items in the context of a physical science assessment. Paper presented at the biannual meeting of the International Objective Measurement Workshop, Philadelphia, PA.
- Wind, S. A.** & Engelhard, G., Jr. (2014, April). Monitoring rating quality in writing assessment using Mokken scale analysis. Paper presented at the annual meeting of the American Educational Research Association, Philadelphia, PA.

- Cheong, Y. F., **Wind, S. A.**, Engelhard, G., Jr., & Kaliski, P. (2013, October). 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 annual meeting of the Northeastern Educational Research Association, Rocky Hill, CT.
- Hendrickson, A., **Wind, S. A.**, Morgan, D., & Kaliski, P. (2013, October). Using multivariate generalizability theory to assess consistency of standard setting ratings. Paper presented at the annual meeting of the Northeastern Educational Research Association, Rocky Hill, CT.
- Wind, S. A.**, Kaliski, P., Morgan, D., & Engelhard, G., Jr. (2013, October). 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 annual meeting of the Northeastern Educational Research Association, Rocky Hill, CT.
- Wind, S. A.**, & Engelhard, G., Jr. (2013, July). 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.
- Wind, S. A.**, & Engelhard, G., Jr. (2013, April). How invariant and accurate are domain ratings in writing assessment? Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Wind, S. A.**, & Engelhard, G., Jr. (2012, November). The influences of panelist and item characteristics on standard-setting judgments within the context of the Advanced Placement world language examinations. Paper presented at the conference to honor Ronald K. Hambleton (“The Ronference”), Amherst, MA. (**Received Graduate Student Award**).
- Engelhard, G., Jr. & **Wind, S. A.** (2012, July). Evaluating rater-mediated assessments. Paper presented at the meeting of the International Test Commission, Amsterdam.
- 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.
- Wind, S. A.**, & Engelhard, G., Jr. (2012, April). Examining gender differences in writing assessment using rating quality indices. Paper presented at the annual meeting of the International Objective Measurement Workshop, Vancouver.
- Engelhard, G., Jr., **Wind, S. A.**, Kobrin, J., & Chajewski, M. (2012, April). 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.
- Engelhard, G. & **Wind, S. A.** (2012, April). Rating quality studies using Rasch measurement theory. Paper presented at the annual meeting of the American Educational Research Association, Vancouver.

- Hendrickson, A., **Wind, S. A.**, & Engelhard, G., Jr., (2012, April). Comparative analyses of generalizability theory and the many-facet Rasch model. Paper presented at the annual meeting of the American Educational Research Association, Vancouver.
- Kaliski, P. K., **Wind, S. A.**, Engelhard, G., Jr., Morgan, D. L., Reshetar, R. A., & Plake, B. L. (2012, April). 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.
- Cheong, Y. F., Engelhard, G., Jr., & **Wind, S. A.**, (2012, April). Hierarchical generalized linear models for rater-mediated assessments. Paper presented at the annual meeting of the American Educational Research Association, Vancouver.
- Wind, S. A.**, & Engelhard, G., Jr. (2011, October). Examining the congruence among rating quality indices in writing assessment: Rater agreement, error, and accuracy. Paper presented at the annual meeting of the Georgia Educational Research Association, Savannah, GA.
- Wind, S. A.**, & Engelhard, G., Jr. (2011, July). Evaluating the quality of ratings in writing assessment: Agreement, errors, and accuracy. Paper presented at the Pacific Rim Objective Measurement Seminar, Singapore. **(Received Graduate Student Research Award).**

INVITED PRESENTATIONS

- Wind, S. A.** (2017, April). Exploring rating scale category formulations across polytomous Rasch and Mokken models. Invited presentation at the business meeting of the Rasch Measurement Special Interest Group at the annual meeting of the American Educational Research Association in San Antonio, Texas.
- Wind, S. A.** (2015, October). Evaluating rater-mediated assessments using Mokken scaling. Invited presentation at the monthly quantitative psychology department brownbag session at the Georgia Institute of Technology, Atlanta, Georgia.

GRANTS

- Subcontract with National Science Foundation**, MSP Division, NSF 1238089, \$19,979 2016 – 2018
Grant title: *Advanced Manufacturing and Prototyping Integrated to Unlock Potential (AMP-IT-UP)*
Position: Data analysis
- National Science Foundation**, MSP Division, NSF 1238089, \$7.9M 2013 – 2015
Grant title: *Advanced Manufacturing and Prototyping Integrated to Unlock Potential (AMP-IT-UP)*
Position: Research support

- National Science Foundation**, DR-K-12 division, NSF 0918618, \$3.5M 2013 – 2015
 Grant title: *Science Learning: Integrating Design, Engineering, and Robotics (SLIDER)*
 Position: Research support
- Pearson**, Human and Automated Scoring, \$10,000 August – December 2014
 Grant title: *Exploring the effectiveness of rater training for teachers.*
 Position: Data analysis and support
- Pearson**, Human and Automated Scoring, \$10,000 August – December 2013
 Grant title: *Examining the influences of rater characteristics on automated scoring engines.*
 Position: Primary Investigator
- College Board**, Office of Research and Development, \$5,000 January – June 2012
 Grant title: *Rating quality studies using Rasch measurement theory*
 Position: Graduate Research Assistant
- College Board**, Office of Research and Development, \$5,000 January – June 2012
 Grant title: *Differential item and person functioning in large-scale writing assessments within the context of the SAT Reasoning Test*
 Position: Graduate Research Assistant

ACADEMIC AWARDS

- Alicia Cascallar Award for an Outstanding Paper by an Early Career Scholar** April 2018
National Council on Measurement in Education
 Presented at the 2018 meeting in New York, NY
- Exemplary Paper Award** April 2017
Classroom Observation Special Interest Group, AERA
 Presented at the 2017 meeting in San Antonio, TX
- Georg William Rasch Early Career Publication Award** April 2015
Rasch Measurement Theory Special Interest Group, AERA
 Presented at the 2015 meeting in Chicago, IL
- Graduate Student Research Award** April 2014
International Objective Measurement Workshop, Philadelphia, PA
- Graduate Student Award** November 2012
Teach Your Children Well: A Conference to Honor Ronald K. Hambleton, Amherst MA
- Division of Educational Studies Full Fellowship** 2012 – 2014
 Emory University, Atlanta, GA

Laney Graduate School Tuition Grant Emory University, Atlanta, GA	2010 – 2014
Graduate Student Research Award <i>Pacific Rim Objective Measurement Symposium</i> , Singapore	July 2011
William S. Pickens Endowment for Education Scholarship University of West Florida, Pensacola, FL	2007 – 2009
Alpha Delta Kappa Education Scholarship University of West Florida, Pensacola, FL	2007 – 2009
Florida Bright Futures Academic Scholars Scholarship University of West Florida, Pensacola, FL	2006 – 2009
Presidential Scholarship University of West Florida, Pensacola, FL	2006 – 2009

EDUCATIONAL TESTING INDUSTRY APPOINTMENTS

Pearson Psychometric Intern Human and Automated Scoring	Summer – Fall 2013
Data Recognition Corporation Psychometric Intern Psychometric Services	Summer 2013
College Board Psychometric Intern Office of Research and Development	Summer 2012
College Board Graduate Research Assistant Office of Research and Development	2011 – 2014

TEACHING EXPERIENCE

The University of Alabama	2015 – present
<ul style="list-style-type: none"> • Fall 2018: <ul style="list-style-type: none"> ○ Quantitative Research Methods (Graduate course) ○ Quantitative Methods II (Graduate course) 	

- **Spring 2018:**
 - Quantitative Methods II (Graduate course)
 - Introduction to Psychometrics (Graduate course)
- **Fall 2017:**
 - Quantitative Research Methods (Graduate course)
 - Quantitative Methods II (Graduate course)
- **Spring 2017:**
 - Quantitative Methods II (Graduate course)
 - Introduction to Psychometrics (Graduate course)
- **Fall 2016:**
 - Introduction to Rasch Measurement Theory (Graduate course)
- **Spring 2016:**
 - Quantitative Methods II (Graduate course)
 - Introduction to Psychometrics (Graduate course)
- **Fall 2015:**
 - Quantitative Methods II (Graduate course)

University of Illinois at Chicago, Chicago, IL 2014 – 2015
 Online module for Educational Measurement (Graduate course): Equating methods for judge-mediated data using Facets

Emory University, Atlanta, GA 2011 – 2015
 Instructor, Division of Educational Studies

- **Spring 2015:** Instructor, Educational Measurement (Undergraduate course)
- **Fall 2012:** Instructor, Educational Measurement (Undergraduate course)
- **Summer 2011:** Co-instructor, Educational Measurement (M.A. in Teaching course)
- **Fall 2011:** Teaching Assistant, Educational Measurement (Undergraduate course)

Emory University, Atlanta, GA 2013 – 2014
 Instructor, Institute for Quantitative Theory and Methods

- **Fall 2013:** Teaching Assistant and Lab Instructor, Introductory Statistics (Undergraduate course)

Green Forest Academic Tutoring, Atlanta, GA 2010 – 2011
 Teacher: Remedial English Language Arts and Mathematics, 6th grade

Edge Elementary School, Niceville, FL 2010
 Teacher: Remedial Writing and Mathematics, 4th grade

Edge Elementary School, Niceville, FL 2009
 Student Teacher, 4th grade

WORKSHOP TEACHING EXPERIENCE

International Objective Measurement Workshop, New York, NY April 2018

Workshop instructor, *Intermediate Winsteps Techniques*. Post-conference software workshop at the bi-annual meeting of the International Objective Measurement Workshop.

Georgia Institute of Technology, Atlanta, GA February 2014
Workshop instructor, *Introduction to measurement: Rasch measurement theory and applications*. Center for Education Integrating Science, Mathematics, and Computing (CEISMC) professional development seminar series.

Data Recognition Corporation, Maple Grove, MN October 2013
Graduate Teaching Assistant, *Introduction to Rasch measurement: Theory and applications*, Journal of Applied Measurement (JAM) Press Workshop.

College Board, New York, NY September 2011
Graduate Teaching Assistant, *Invariant rater-mediated measurement: Using the Facets computer program to evaluate the quality of rater-mediated assessments with Many-Facet Rasch Models*. Office of Research and Development Workshop.

Pacific Rim Objective Measurement Seminar, Singapore July 2011
Graduate Teaching Assistant, *Using Rasch models to evaluate the quality of performance assessments*. Pre-conference workshop at the Pacific Rim Objective Measurement Seminar, Singapore.

SERVICE TO PROFESSIONAL COMMUNITY

Program Chair, *Classroom Observation Special Interest Group, American Educational Research Association* 2018 – present

Co-Editor, *Rasch Measurement Transactions* 2018 – present

External Dissertation Reviewer, *The University of Adelaide, Australia* Summer 2016

Editorial Board Member, *Measurement: Interdisciplinary Research and Perspectives* 2017 – present

Editorial Board Member, *Journal of Educational Measurement* 2016 – present

Conference planning committee *International Objective Measurement Workshop* 2015 – present

Reviewer for academic journals:

- *Applied Psychological Measurement*
- *Assessing Writing*
- *Educational and Psychological Measurement*
- *Educational Measurement: Issues and Practice*
- *Language Testing*
- *Journal of Geoscience Education*

- *Measurement: Interdisciplinary Issues and Practice*
- *Musicae Scientiae*
- *Statistics Education Research Journal*
- *Sociological Methodology*

SERVICE TO UNIVERSITY

Quantitative Research Certificate Co-Coordinator Department of Educational Research, The University of Alabama	Fall 2018 – Spring 2019
Jim McClean Graduate Student Dissertation Award committee member , Department of Educational Research, The University of Alabama	Fall 2016
Search committee member Department of Educational Research, The University of Alabama	Fall 2016
Measurement curriculum planning committee Department of Educational Research, The University of Alabama	Fall 2015 – present
Undergraduate Honors Thesis Committee Member (reader) Emory University	2014 – 2015
Statistics Lab Administrator Emory University	2011 – 2014
Graduate Student Research Roundtable Coordinator Emory University	2010 – 2014

PROFESSIONAL ORGANIZATION MEMBERSHIPS

American Educational Research Association

Division D; Classroom Observation Special Interest Group; Rasch Measurement Theory Special Interest Group; Science Teaching and Learning Special Interest Group

American Harp Society

Georgia Educational Research Association

International Test Commission

National Council on Measurement in Education