**VITA**

**STONE, Clement A.**

**(Spring 2017)**

**PERSONAL**

Office Address:

5920 WWPH

University of Pittsburgh

Pittsburgh, PA 15260

412-624-9359 email: cas@pitt.edu

ACADEMIC BACKGROUND

B.S. 1977 St. Lawrence University, Canton, NY; Major: Psychology

M.Ed. 1980 University of Arizona, Tucson, AZ; Major: Educational Psychology

Ph.D. 1987 Department of Educational Psychology, University of Arizona, Tucson, AZ

Major: Research Methodology, Measurement, and Statistics

Minor: Management Information Systems; Cognates: Mathematics, Computer Science

Dissertation: On the Robustness of Total Indirect Effects in the Jöreskog-Keesling-Wiley Covariance Structure Model.

# AWARDS

Outstanding Reviewer Award – American Educational Research Journal - 2002

College of Education Graduate Student Research Award – 1986

Award of Excellence – Society for Technical Communication, Carolina Chapter – 2015-2016

RESEARCH INTERESTS

Research interests include educational and psychological assessment, large scale assessment and psychometrics, including instrument development and validation, item response theory models and applications, computer adaptive testing, simulation and Monte Carlo methods, applied Bayesian methods and applications to IRT, quasi-experimental designs and propensity scoring methods, and the measurement of change.

**EMPLOYMENT HISTORY**

2006-present Professor, Psychology in Education, School of Education, University of Pittsburgh

2002-2006 Associate Professor, Psychology in Education, School of Education, University of Pittsburgh

1998-2002 Assistant Professor, Psychology in Education, School of Education, University of Pittsburgh

1996-1998 Assistant Professor, Center for Nursing Research, School of Nursing, University of Pittsburgh

1987-1996 Research Specialist, CIS, University of Pittsburgh

1991-1993 Psychometric Consultant to QUASAR Project, LRDC, University of Pittsburgh

1981-87 Research Staff Member, Measurement Center, University of Arizona

**PROFESSIONAL ORGANIZATION MEMBERSHIPS**

Member of American Educational Research Association (AERA)

Member of National Council on Measurement in Education (NCME)

Member of American Psychological Association (Division 5)

**PROFESSIONAL ACTIVITIES**

Member, NCME Awards Committee (2007-2010)

Chair, AERA Division D Nominating Committee (2001-2002)

Member, NCME Dissertation Award Committee (2000-2001; 2013-2014)

Member, AERA Division D Nominating Committee (2000-2001, 2002-2003)

Proposal Reviewer, AERA Annual Meeting, Division D

Proposal Reviewer, NCME Annual Meeting

Proposal Reviewer, NAEP 2002 Secondary Analysis Grant Program (NCES)

Proposal Reviewer, Center for Research in Chronic Disorders, University of Pittsburgh School of Nursing

Proposal Reviewer, Western Psychiatric Institute and Clinic, University of Pittsburgh

Psychometric Consultant to Pennsylvania Alternate Science Assessment (PASA) (2015-present)

Member, Statistics and Methodology Core, University of Pittsburgh – School of Pharmacy (2003-2015)

Psychometric Consultant to United Steel Workers (2008-2009)

Psychometric Consultant for the DRA Reading Assessment Program, Pearson Learning Group (2003-2007)

Assessment Consultant, Federal Express Corporation (2006)

Statistical Consultant to West Allegheny School District (2001-2002), Pittsburgh PA

Psychometric Consultant for the Competencies Assessment Program, DeVry Inc., Chicago (1986-1996)

Member, Panel of Experts, Indiana Performance Assessment for Students' Success (IPASS) Program (1995)

EDITORIAL ACTIVITIES

Editorial Board, Journal of Educational Measurement (2011 to 2017)

Editorial Board, Applied Measurement in Education (2009 to 2017)

Editorial Board, Educational and Psychological Measurement (2008 to 2014)

Editorial Board, American Educational Research Journal (2001 to 2004)

Reviewer, Journal of Educational Measurement

Reviewer, Applied Psychological Measurement

Reviewer, Educational Measurement Issues and Practice

Reviewer, Sociological Methodology Research

Reviewer, American Educational Research Journal

Reviewer, Review of Educational Research

Reviewer, Psychometrika

Reviewer, Psychological Methods

Reviewer, Comparative Education Review

Reviewer, International Journal of Testing

Reviewer, Journal of Speech, Language, and Hearing Research

TEACHING EXPERIENCE

* 1. University of Pittsburgh, School of Education: Graduate Level.

Factor Analysis and Instrument Validation, (3 Credits)

Constructing and Analyzing Surveys, (3 Credits)

Advanced Computer Applications in Educational Research – Introduction to Monte Carlo Experimentation, and Applied Bayesian Analysis (3 Credits)

Item Response Theory and Applications (3 Credits)

Educational Testing and Measurement (3 Credits)

Constructing Achievement and Ability Tests (3 Credits)

Experimental and Quasi-experimental Design (3 Credits)

**GRANT AWARDS**

2014-2016 Stone (Co-Investigator; PI: T. Hagan). *Development and Validation of a Measure of Self-Advocacy in Female Cancer Survivors*. American Cancer Society Doctoral Student in Nursing Research Award.

2012-2014 Stone (Co-Investigator; PI: S. Lane). *Research on the Effectiveness of a remote coaching model*, Bill and Melinda Gates Foundation.

2012-2013 Stone (Co-Investigator; PI: S. Lane). *Research/Evaluation of the Teacher Evaluation Project for the Pennsylvania Department of Education*, PDE and US Department of Ed *Race to the Top Funding*.

2011-2013 Stone (Co-Investigator; T. Kamarck, PI). *Novel measures of psychosocial stress*, National Institute of Drug Abuse.

2010-2011 Stone (Co-Investigator; PI: S. Lane). *Research/Evaluation Work Plan for a Teacher and Principal Evaluation Project for the Pennsylvania* *Department of Education,* Team PA Foundation and Bill and Melinda Gates Foundation

7/10 – 6/15 Hula, W. (Mentor: Stone, C.). *Communication Outcome Measurement in Aphasia.* VA Rehabilitation Research and Development Level 2 Career Development Award, Department of Veteran Affairs (# C7476W).

6/10- 9/11 Stone (Co-Investigator: Psychometrician, Statistician; PI: C. Tananis). *Evaluating the effect of the MSP intervention on teachers and student achievement*. Evaluation services for Pittsburgh Public Schools; Funded by Commonwealth of PA and Allegheny Intermediate Unit.

9/08-11/10 Stone (Co-Investigator: Psychometrician, Statistician; PI: C. Tananis). *Math Science Partnership of Southwest Pennsylvania: Measuring Progress Towards Goals*, NSF: EH4 0314914.

7/08 – 6/11 Stone (Co-Investigator; P. Doyle, PI). *Communicative Functioning in Neurogenic Speech/Language Disorders II*, Veterans Affairs Rehabilitation Research and Development Service – 151U-H.

5/08 – 6/10 Stone (Co-Investigator; T. Kamarck, PI). *Psychosocial stress exposure: Real time and structured interview technologies*, National Institute of Drug Abuse - Award DA023821.

5/07 – 9/07 Zhu, X. (Mentor: Stone). *An Evaluation of Different Approaches to Subscore Augmentation for the Multistate Bar Examination*, funded by the Multistate Bar Association.

6/06 – 11/06 Stone, C.A. (PI). *Augmenting subscale scores for the Delaware State Assessment Program. State Department of Delaware*, Delaware State Department of Education.

1/06 – 9/06 Stone, C.A. (PI). *Evaluating the DRA2 Reading Assessment*, Pearson Learning Group.

5/06 – 9/06 Stone, C.A. (PI*). Evaluating the ACTS of PA Curriculum Intervention: A Meta-Analysis using a Quasi-experimental Design*. SERVE, University of North Carolina.

5/05 – 12/05 Stone, C.A. (PI). *Evaluating the ACTS of PA Curriculum Intervention: Implementing Propensity Scoring in a Quasi-Experimental Design*, SERVE, University of North Carolina.

6/03 – 12/03 Yen, C. (Mentor: Stone). *Assessing the dimensionality and factor structure for the Multistate Bar Exam*, funded by the Multistate Bar Association

8/01 - 7/02 Stone, C.A. (PI). *Simulating computer adaptive tests*, University of Pittsburgh, Central Research Development Fund.

###### 6/01-7/02 Dresher, A. (Mentor: Stone). *The examination of local item dependency of NAEP assessments using the testlet model*. 2001 Harold Gulliksen Psychometric Fellowship Program, Educational Testing Service.

8/01-7/04 Stone, C.A. (Co-investigator). *Assessing the consequences of the Pennsylvania State Performance Assessment Program (S. Lane, PI)*, Commonwealth of Pennsylvania, State Department of Education .

7/00 - 7/01 Stone, C.A. (PI). *Evaluation of a Wald Test for Assessing the Significance of a Fit Statistic Based on Posterior Expectations in Item Response Theory Models*, Department of Education (Grant # R902B000008).

12/95-6/00 Stone, C.A. (Co-investigator). *Assessing the consequences of the Maryland State Performance Assessment Program (S. Lane, PI)*, Maryland State Dept. of Education.

1/98-10/99 Stone, C.A. (PI). *Assessing Goodness of Fit of IRT Models to NAEP Data*, Department of Education (Grant # R902B970008).

9/97-6/99 Stone, C.A. (PI). *Using a General Measure of Critical Thinking to Measure Nursing Students’ Critical Thinking*, University of Pittsburgh, Central Research Development Fund.

**BOOKS and CHAPTERS**

Stone, C.A. & Zhu, X. (2015). *Bayesian analysis of item response theory models using SAS*. Cary, NC: SAS Publications.

Lane, S. & Stone, C.A. (2006). Performance Assessment. Chapter in R. L. Brennan (Ed.),  *Educational measurement* (4th ed.).  Westport, CT: American Council on Education/Praeger.

**PUBLICATIONS (Measurement and Statistics)**

Kamarck, T.W., Absi, M., Epstein, D. Estin, E., Intille, S., Kirk, G. Kumar, S. Preston, K. Rea, M,. Scott, M., Shetty, V., Shiffman, S,. Stewiorek, D., Smailagic, A. Stone, C.A., & Venugapa, M. (2017). Amabulatory Monitoring and Ecological Momentary Assessment. In Waldstien, S. Katzel, L. & Kop, W. (Eds.), C*ardiovascular behavioral medicine*.

Hagan, T. L., Cohen, S.M., Stone, C.A., & Donovan, H. D. (2016). Theoretical to tangible: Creating a measure of self-advocacy for female cancer survivors. *Journal of Nursing Measurement*. 24 (3), 428-441.

Stone, C.A., & Leventhal B.C. (2016). Accounting for multidimensionality in item responses in patient-centered and patient reported outcomes measurement. *International Journal of Person Centered Medicine*, 6(4), 274-288.

Hula, W.D., Doyle, P.J., Stone, C.A., Hula, S.N., Kellough, S., Wambaugh, J.L., Ross, K.B., Schumacher, J.G., & Jacque, A.S. (2015). The Aphasia Communication Outcome Measure (ACOM): Dimensionality, item bank calibration, and initial validation. *Journal of Speech, Language, and Hearing Research,* 58(3), 906-916.

Doyle, P.J., Hula, W.D., Hula, S.N., Stone, C.A., Wambaugh, J.L., Ross, K.B., & Schumacher, J.G. (2013). Self- and surrogate-reported communication functioning in aphasia. *Quality of Life Research*, 22(5), 957-67.

Stone, C. A. & Tang, Y. (2013). Comparing propensity score methods in balancing covariates and recovering impact in small sample educational program evaluations. *Practical Assessment, Research & Evaluation*, 18(13). Available online: <http://pareonline.net/getvn.asp?v=18&n=13>

Stone, C.A. (2013). Using SAS PROC MCMC to estimate and evaluate item response theory models. Proceedings from SIS 2013 published in ***Advances in Latent Variables***, Eds. Brentari E., Carpita M., Vita e Pensiero, <http://meetings.sis-statistica.org/index.php/sis2013/ALV/paper/view/2541>.

Kirisci, L., Tarter, R., Reynolds, M., Ridenour, T., Stone, C., & Vanyukov, M. (2012). Computer adaptive testing of liability to addiction: Identifying individuals at risk. *Drug and Alcohol Dependence*. Suppl 1:S79-86.

Judith A. C., Dunbar-Jacob, J., Sereika, S.M., Stone, C.A., Fasickzka, A., Jarrett, R.B., & Thase, M.E. (2012). Barriers to Cognitive Behavioral Therapy Homework Completion Scale- Depression Version: Development and Psychometric Evaluation. *International Journal of Cognitive Therapy*, 5 (2), 219-235.

Zhu, X., & Stone, C.A. (2012). Bayesian comparison of alternative graded response models for performance assessment applications. *Educational and Psychological Measurement*. *72*,774-799*.*

Xu, T., & Stone, C.A. (2012). Using IRT trait estimates versus summated scores in predicting outcomes. *Educational and Psychological Measurement*, *7*, 453-468.

Zhu, X., & Stone, C.A. (2011). Assessing fit of unidimensional graded response models using Bayesian methods. *Journal of Educational Measurement*, *48*, 81-97.

Stone, C.A., Ye, F., Zhu, X., & Lane, S. (2010). Providing subscale scores for diagnostic information: A case study when a test is essentially unidimensional. *Applied Measurement in Education*, 23, 63-86.

Zhang, B. & Stone, C.A. (2008). Evaluating item fit for compensatory multidimensional item response theory models. *Educational and Psychological Measurement*, *68* (2), 181-196.

Bjorner, J.B, Smith, K.J., Stone, C.A., & Sun, X. (2007). *IRTFIT: A macro for item fit and local dependence tests under IRT models*. Lincoln, RI: Quality Metric, Inc.

Stone, C.A., & Yeh, C. (2006). Assessing the dimensionality and factor structure of multiple-choice exams: An empirical comparison of methods using the Multistate Bar Examination. *Educational and Psychological Measurement*, *66,* 193-214.

Parke, C. S., Lane, S., & Stone, C. A. (2006). Impact of a state performance assessment program in reading and writing. *Educational Research and Evaluation*, *12* (3), 239-269.

Stone, C.A., Weissman, A., & Lane, S. (2005). Consistency of student proficiency classifications under competing IRT models for a state assessment program. *Educational Assessment*, *10*, 125-146.

Irrgang, J. J. and Stone, C. A. (2004). Simulation of a computer-adaptive test utilizing the disabilities of the arm, shoulder, and hand (DASH) outcomes measure. *Journal of Orthopaedic and Sports Physical*, Therapy, *34*(1), A15.

Stone, C.A., Irrgang, J.J. (2004). Simulation of a computer-adaptive test utilizing the Disabilities of the Arm Shoulder Hand (DASH) outcome measure. *Journal of Orthopaedic and Sports Physical Therapy*, 34(1), A15.

Stone, C.A. (2004) IRTFIT Resample: A computer program for assessing goodness-of-fit of IRT models when ability is measured imprecisely. *Applied Psychological Measurement*, 28, 143-145.

Stone, C.A. & Zhang, B. (2003). Comparing three new approaches for assessing goodness-of-fit in IRT models. *Journal of Educational Measurement*, *4*, 331-352.

Stone, C.A. (2003). Empirical power and Type I error rates for a goodness-of-fit statistic based on posterior expectations and resampling-based inference. *Educational and Psychological Measurement, 63*, 566-583.

Stone, C.A. & Lane, S. (2003). Consequences of a State Accountability Program: Examining Relationships between School Performance Gains and Teacher, Student, and School Variables. *Applied Measurement in Education*, *1,* 1-26.

Lane, S., Parke, C.S., & Stone, C.A.  (2002). The Impact of a State Performance-Based  Assessment and Accountability Program on Mathematics Instruction and Student Learning: Evidence from Survey Data and School Performance. *Educational Assessment*, 8, 279-316.

Lane, S. & Stone, C.A. (2002). Strategies for examining the consequences of assessment and accountability programs. *Educational Measurement: Issues and Practice, 21*, 23-30.

Stone, C.A., Davidson, L. Evans, J.L, & Hansen, M.A. (2001). Validity evidence for using a general critical thinking test to measure nursing students’ critical thinking. *Holistic Nursing Practice, 15*, 65-74

Stone, C.A & Hansen, M.A. (2000). The effect of errors in estimating ability on goodness of fit tests for IRT models. *Educational and Psychological Measurement, 60*, 974-991

Stone, C.A & Hansen, M.A. (2000). Uncertainty in ability estimation and assessing goodness-of-fit of IRT models. *Proceedings of the International Conference on Measurement and Multivariate Analysis, 2*, 194-192.

Stone, C.A. (2000). Monte-carlo based null distribution for an alternative fit statistic. *Journal of Educational Measurement, 37,* 58-75.

Lane, S., Parke, C.S., & Stone, C.A. (1998). A framework for evaluating the consequences of assessment programs. *Educational Measurement: Issues and Practice, 17*, 24-28.

Harwell, M., Stone, C.A., Hsu, T.C., & Kirisci, L. (1996). Monte Carlo studies in item response theory. *Applied Psychological Measurement, 20*, 101-126.

Lane, S., Liu, M., Ankenmann, R.D., & Stone, C.A. (1996). Validity and generalizability of a performance assessment. *Journal of Educational Measurement, 33*, 71-92.

Lane, S., Stone, C.A., Ankenmann, R.D., & Liu, M. (1995). Examination of the assumptions and properties of the graded item response model: An example using a mathematics performance assessment. *Applied Measurement in Education, 8,* 313-340.

Lane, S., Stone, C.A., Ankenmann, R.D., & Liu, M. (1994). Reliability and validity of a mathematics performance assessment. *International Journal of Educational Research, 21*, 247-262.

Stone, C.A. (1993) Book Review: Using national databases in educational research. *Educational Measurement: Issues and Practice, 12,* 37-39.

Stone, C.A. (1992). TESTAT: A supplementary module to SYSTAT. *Educational Measurement: Issues and Practice, 11,* 33.

Stone, C.A. (1992). Recovery of marginal maximum likelihood estimates in the two parameter logistic response model: An evaluation of MULTILOG. *Applied Psychological Measurement, 16*, 1-16.

Stone, C.A. & Lane, S. (1991). Use of restricted item response theory models for examining the stability of item parameter estimates over time. *Applied Measurement in Education, 4*, 125-142.

Stone, C.A. (1990). Software Review - STAT Version 1.0. *Educational Measurement: Issues and Practice, 9*, 27-30.

Stone, C.A. & Sobel, M.E. (1990). The robustness of estimates of total indirect effects in covariance structure models estimated by maximum likelihood. *Psychometrika, 55,* 337-352.

Stone, C.A. (1989). Microcomputer applications to testing: A review of MicroCAT version 3.0. *Educational Measurement: Issues and Practice, 8*, 33-38.

Bergan, J.R., & Stone, C.A. (1986).  Psychometric and instructional validation of hierarchical domain structures.  *Contemporary Educational Psychology, 11*, 1-32.

Feld, J.K., Bergan, J.R., & Stone, C.A. (1986).  Behavioral consultation:  Current status, future directions.  In C.A. Maher & S.G. Forman (Eds.), *Providing effective educational services in school organizations: A behavioral approach*.  Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bergan, J.R., Stone, C.A., & Feld, J.K. (1985). Path‑referenced evalua­tion of individual differences.  In C.R. Reynolds and Wilson, V.L. (Eds.) *Methodological and Statistical Advances in the Study of Individual Differences.* New York: Plenum.

Stone, C.A.  (1985).  CINDESE: Computing indirect effects and their standard errors.  *Educational Psychology and Measurement, 45*, 601-606.

Bergan, J.R., & Stone, C.A. (1985).  Latent class models for knowledge domains. *Psychological Bulletin, 98*, 166‑184.

Bergan, J.R., Stone, C.A., & Feld, J.K.  (1984).  Rule replacement in the development of basic number skills.  *Journal of Educational Psychology, 76*, 289‑299.

Stone, C.A., & Bergan, J.R.  (1983).  Measuring cognitive growth in preschool education programs.  *New Jersey Journal of School Psychology, 2*, 76‑86.

**OTHER PUBLICATIONS (Health Related Empirical Studies)**

Constantino, R.E., Sekula, L.K., Rabin, B., & Stone, C.A. (2000). Negative life experiences, depression, and immune function in abused and nonabused women. *Biological Research For Nursing, 1*, 190-198.

Albrecht, S.A., Higgins, L.W., & Stone, C.A. (1999). Factors relating to pregnant adolescents’ decisions to complete a smoking cessation intervention. *Journal of Pediatric Nursing, 14*, 322-327.

Albrecht, S.A., Cornelius, M.D., Braxter, B., Reynolds, M.D., Stone, C.A., & Cassidy, B. (1999). An assessment of nicotine dependence among pregnant adolescents. *Journal of Substance Abuse Treatment, 16,* 337-344.

Albrecht, S.A., Payne, L.J., Stone, C.A., & Reynolds, M.R. (1998). A preliminary study of the use of peer support in smoking cessation interventions for pregnant adolescents. *Journal of the American Academy of Nurse Practitioners, 10*, 119-125.

Whitman, G.R., Weber, M.M., Stovsky, B., Sereika, S.M., Stone, C.A., Blackburn, G.G., & Stoller, J.K. (1998). Predictors of post hospital discharge costs at one and six months after cardiac surgery. *American Heart Association 71st Scientific Sessions, Circulation Supplement I, 98*, 82.

Grote, N.K., Frieze, I.H., & Stone, C.A. (1996). Children, traditionalism in the division of family work, and marital satisfaction: “What’s love got to do with it?” *Personal Relationships, 3*, 211-228.

**PRESENTATIONS**

Levanthal, B.C., & Stone, C.A. (April, 2017). Modeling extreme response style. Paper presented at the Annual Meeting of the National Council for Measurement in Education, San Antonio, Texas.

Stone, C.A. (June, 2016). *Bayesian analysis using SAS PROC MCMC*. Invited workshop (18 credits), School of Government and Public Policy, Beijing Normal University.

Hagan, T. L., Cohen, S. M., Rosenzweig, M. Q., Zorn, K. K., Stone, C. A., Donovan, H. S. (January, 2016). *The Female Self-Advocacy in Cancer Survivorship Scale: Validating a new measure to support patient voice and engagement*. Abstract accepted for the 2016 Cancer Survivorship Symposium: Advancing Care and Research, San Francisco, CA.

Iwatani, E., Stone, C.A., Shealy, C.N. (April, 2016). *How classification trees can be helpful to educational research: A case study.* Paper presented at the Annual Meeting of the American Educational Research Association, Division D, Washington, D.C.

Levanthal, B.C., & Stone, C.A. (April, 2016). Extreme response style: Which model is best? Paper presented at the Annual Meeting of the National Council for Measurement in Education, Washington, D.C.

### Scott, P.W., Stone, C.A., Bachman, H.J., Elliot L. (April, 2016). *Comparing Heckit and Propensity Methods Under Conditions When There Is Evidence of Selection Into Treatment.* Paper presented at the Annual Meeting of the American Educational Research Association, Division D, Washington, D.C.

Stone, C.A. (April, 2016). *Bayesian analysis of item response theory models using SAS*. Workshop at the Annual Meeting of the National Council for Measurement in Education, Washington, D.C.

Leventhal, B., Stone, C.A., Yu, L., & Greco, C. (April, 2015). *Comparing MULTILOG and IRTPRO parameter estimates under skewed trait distributions*. Paper presented at the Annual Meeting of the National Council for Measurement in Education, Chicago.

Stone, C.A. (April, 2014). *Using propensity scoring methods in program evaluation research*. Presentation at the Eastern Evaluation Research Society Conference, Galloway, NJ.

Stone, C.A. (June, 2013). *Using SAS PROC MCMC to estimate and evaluate item response theory models*. Paper presented at the SIS annual meeting, Brescia, Italy.

Kamarck, T.W., Anderson, B., Siewiorek, D.P., Stone, C.A., Shiffman, S., Smailagic, A.,  
Wethington, A., French, B. , Heineman, E. (March, 2013). *Novel measurement methods for assessing psychosocial stress: Preliminary validation*. Paper presented at the annual meeting of the American Psychosomatic Society, Miami.

Doyle, P. Hula, W., Hula, S., Stone, C.A., Ross, K., Wambaugh, J., Schumacher, J. & Roehrig, A. (May, 2012). *The aphasia communication outcome measure: Item reduction, scaling, and concurrent validity of self-reported communicative functioning in effects of modality and word frequency in sentence comprehension in people with aphasia*. Paper presented at the **Annual Clinical Aphasiology Conference, Lake Tahoe, NV.**

Stone, C.A., & Tang, Y. (April, 2012). *Comparing propensity scoring methods: Practical issues for studying program impact*. Paper presented at the 2012 Eastern Evaluation Research Society Conference, Galloway, New Jersey.

Stone, C.A. (February, 2012). *Reliability and validity evidence in survey research*. Invited presentation for the MERMAID lecture series. University of Pittsburgh Medical Education Department.

Tang, Y., & Stone, C.A. (April, 2011). *Evaluating student achievement from large scale assessments: A case study using propensity score methods*. Paper presented at the Annual Meeting of the National Council for Measurement in Education, New Orleans.

Kamarck, T.W., Siewiorek, D., Smailagic A., Shiffman S., Stone C.A., French B., Beach S., Heineman E., Lubinski L., Anderson B. (November, 2010). *Computer adaptive testing methods for EMA assessment of psychosocial stress*. Technology Demonstration Presentation, 2010 mHealth Summit, Washington, DC.

Stone, C.A. (November, 2010). *Surveys: Developing and designing survey research*. Invited presentation for the MERMAID lecture series. University of Pittsburgh Medical Education Department.

Tang, Y., & Stone, C.A. (April, 2010). *Comparing Different Propensity Scoring Methods in a Quasi-Experimental Design Using Student Achievement Results*. Paper presented at the Annual Meeting of the American Educational Research Association, Division D, Denver, CO.

Xu, T., & Stone, C.A. (April, 2010). *Using IRT trait estimates versus summated scores in predicting outcomes.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, Denver, CO.

Zhu, X., & Stone, C.A. (April, 2010). *Bayesian comparison of alternative graded response models for performance assessment applications*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Denver, CO.

Kamarck TW, Siewiorek D, Smailagic A, Shiffman S, Stone C, French B, Heineman E, Lubinski L, von Dehsen L, Wolf J, Anderson B, Seigerwalt T. (March 2010). *Developing EMA-based self-report assessment of psychosocial stress*. Presented at Annual meeting of the American Psychosomatic Society, Portland, OR.

Liqun, Y. & Stone, C.A. (2009, October). *Comparisons of IRT Based Ability Estimation Methods and Calibration Procedures in Scaling Test Design*. Paper presented at the Annual Conference of Northeast Educational Research Association, Rocky Hill, Connecticut.

Kirisci, L., Tarter, R., Vanyukov, M., Reynolds, M., Stone, C., & Mezzich, A. (2009, June*). Efficiency and sensitivity of computerized adaptive testing of risk for SUD*. Paper presented at the annual meeting of the College on Problems of Drug Dependence, Scottsdale, AZ.

Zhu,X., & Stone, C.A. (2009, April). *Assessing the fit of the unidimensional GRM model using PPMC methods.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Diego.

Zhu,X., & Stone, C.A. (2009, April). *Detecting local dependence of performance-based items using Bayesian analysis*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Diego.

Stone, C.A., & Lane, S. (2008, June*). Issues in providing subscale scores for diagnostic information*. Paper presented at the National Conference on Student Assessment (CCSSO), Orlando.

Stone, C.A., Tang, Y., & Tierney, S. (March, 2008). *Providing subscale scores for diagnostic information: An evaluation of a regression-based method for augmenting score reliability*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New York.

Zhu, X. & Stone, C.A. (March, 2008). *An evaluation of different approaches to subscore augmentation for the Multistate Bar Examination.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, New York.

Callan, J.A., Dunbar-Jacob, J., Thase, M.E., Jarrett, R., Stone, C.A., Fasiczka, A. (2007, November). *The barriers to CBT Homework Completion Scale: Development and psychometric properties*. Poster session presented at the annual meeting of the Association of Behavioral and Cognitive Therapy, Philadelphia, PA.

Liu, J., & Stone, C.A. (April, 2007). *Comparing multi-dimensional and unidimensional computer-adaptive strategies in psychological and health assessment*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago.

Ye, F., Stone, C., & Lane, S. (April, 2007). *Providing Subscale Scores for Diagnostic Information: A case study of an 8th grade math assessment*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago.

Lane, S., & Stone, C.A. (April, 2007). *New directions in performance assessment*. Invited symposium at the Annual Meeting of the National Council on Measurement in Education, Chicago.

Stone, C.A., & Zhu, X. (November, 2006). *The student achievement study component of the aligning curriculum to standards in Pennsylvania (ACTS of PA) teacher quality enhancement multisite evaluation: A demonstration of a propensity score model*. Paper presented at the American Evaluation Association Conference, Evaluation 2006: The Consequences of Evaluation, Portland, Oregon.

Stone, C.A., & Zhu, X. (August, 2006). *Aligning curriculum to standards (ACTS) in PA: An evaluation using a quasi-experimental design*. Paper presented at the PA Academy for the Profession of Teaching and Learning (2006 Teacher Education Forum), Harrisburg, PA.

Stone, C.A., & Yin, Y. (April, 2006). *Accounting for item and ability parameter estimation in a goodness-of-fit statistic for IRT models.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco.

Stone, C.A., & Irrgang, J.J. (June, 2004). *Simulation of a computer adaptive test utilizing the disabilities of the arm, shoulder and hand (DASH) outcome measure*.  Presented at the DIA Conference “Advances in Health Outcomes Measurement”, Bethesda, MD.

Irrgang, J.J., & Stone, C.A. (June, 2004). *Differential item function for the disabilities of the arm shoulder and hand (DASH) outcome measure*.  Presented at the DIA Conference “Advances in Health Outcomes Measurement”, Bethesda, MD.

Yeh, C. & Stone, C.A. (April, 2004). *Assessing the Dimensionality and Factor Structure of Multiple-Choie Exams: An Empirical Comparison of Methods Using the Multistate Bar Examination*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Diego.

Zhang, B. & Stone, C.A. (April, 2004). *Direct and indirect estimation of a three parameter compensatory multidimensional item response models*. Paper presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego.

Zhang, B. & Stone, C.A. (April, 2004). *Evaluating the sampling distribution of a goodness-of-fit statistic for multidimensional IRT models using total scores*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Diego.

Irrgang J.J., & Stone C.A. (Feb, 2004). *Differential item function for the disabilities of the arm shoulder and hand (DASH) outcome measure*.  Presented at the Combined Sections Meeting of the American Physical Therapy Association, Nashville TN.

Stone, C.A., & Irrgang J.J. (Feb, 2004). *Simulation of a computer adaptive test utilizing the disabilities of the arm, shoulder and hand (DASH) outcome measure*.  Presented at the Combined Sections Meeting of the American Physical Therapy Association, Nashville TN.

Zhang,B. & Stone, C.A. (April, 2003). *Evaluating the sampling distribution of a goodness-of-fit statistic for multidimensional IRT models using total scores*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago.

Stone, C.A. & Zhang, B. (April, 2002). *Comparing three new approaches for assessing goodness-of-fit in IRT models*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans.

Dresher, A.R, Stone, C.A., & Hombo, C. (April, 2002). *The examination of local item dependency of NAEP assessments using the testlet model.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans.

Stone, C.A., Weissman, A., & Lane, S. (April, 2001). *Consistency of student proficiency classifications under competing IRT models for a state assessment program*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Seattle.

Stone, C.A. & Dresher, A. R. (April, 2001). *Using a wald test to evaluate the significance of a goodness-of-fit statistic based on posterior expectations in IRT models*. Paper presented at the 2001 Annual Meeting of the American Educational Research Association, Seattle.

Lane, S. & Stone, C.A. (April, 2001). *Consequences of assessment and accountability programs*. Paper presented at the 2001 Annual Meeting of the American Educational Research Association, Seattle.

Tay-Lim, B.S. & Stone, C.A. (July, 2000). *Assessing the dimensional structure of constructed-response tests using Multidimensional Scaling, Hierarchical Cluster analysis, and Exploratory Factor Analysis*. Paper presented at the Psychometric Society, Vancouver, Canada.

Stone, C.A & Hansen, M.A. (May, 2000). *Uncertainty in ability estimation and assessing goodness-of-fit of IRT models*. Paper presented at the 2000 International Conference on Measurement and Multivariate Analysis, Banff, Alberta, Canada.

Stone, C.A. (April, 2000). *Using resampling methods to evaluate the significance of a goodness-of-fit statistic in item response theory models*. Paper presented at the 2000 Annual Meeting of the American Educational Research Association, New Orleans.

Stone, C.A. & Lane, S. (April, 2000). *MSPAP performance gains from 1993-1998 and their relationship to “MSPAP impact” and school characteristic variables*. Paper presented at the 2000 Annual Meeting of the American Educational Research Association, New Orleans.

Lane, S., Stone, C.A., Parke, C.S., Hansen, M.A., & Cerrillo, T.L. (April, 2000). *Consequential evidence for MSPAP from teacher, principal, and student perspectives (Mathematics, Reading, Writing, Science, and Social Studies).* Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans.

Tay-Lim, B.S. & Stone, C.A. (April, 2000). *Assessing the dimensionality of constructed-response tests using hierarchical cluster analysis: A Monte Carlo study with application*. Paper presented at the 2000 Annual Meeting of the American Educational Research Association, New Orleans.

Stone, C.A. & Lane, S. (February, 2000). *Exploring the impact of a statewide performance assessment program on students, teachers, and schools.* Paper presented at the 2000 Annual Meeting of the Hawaii Educational Research Association, Honolulu.

Lane, S., Cerrillo, T.L., Ventrice, J., Parke, C.S., & Stone, C.A. (April, 1999). *Impact of MSPAP: Evidence from principal, teacher, and student questionnaires*. Paper presented at the National Council for Measurement in Education, Montreal.

Stone, C.A. & Lane, S. (April, 1999). *MSPAP performance gains from 1993-1998 and their relationship to “MSPAP impact” and school characteristic variables*. Paper presented at the National Council for Measurement in Education, Montreal, April.

Riley, T.A. & Stone, C.A. (October, 1999). *The structural relationships of stress, coping self-care, and health outcomes.* Paper presented at the 1999 National Conference on Women and HIV/AIDS, Los Angeles.

Riley, R.A. & Stone, C.A. (February, 1999). *A preliminary structural analysis of stress and health outcomes in women*. Paper presented at the 20th Annual Scientific Sessions of the Society of Behavioral Medicine, San Diego.

Riley, R.A. & Stone, C.A. (April, 1999). *A structural analysis of stress and health outcomes in HIV+ women*. Paper presented at the 23rd Annual Research Conference of the Midwest Nursing Research Society, Indianapolis.

Wang, N., Stone, C.A., & Lane, S. (April, 1998). *An evaluation of the conditional plausible values approach using the graded response model*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Lane, S., Parke, C.S., & Stone, C.A. (April, 1998). *Consequences of the Maryland school performance assessment program*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Albrecht, S.A., Braxter, B., Stone, C. (March, 1998). *Measuring nicotine dependency in pregnant teens who smoke*. Society for Research on Nicotine and Tobacco, 4th Annual Scientific Session, New Orleans.

Lane, S. and Stone, C.A. (April, 1997). *Investigating the consequential aspects of validity: Who is responsible and what should they do*? Invited Symposium at the annual meeting of the American Educational Research Association, Chicago.

Ankenmann, R.D., Stone, C.A., & Lane, S. (April, 1995). *Goodness of fit and ability estimation in the graded response model.* Invited symposia at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

Stone, C.A., Mislevy, R.J., & Mazzeo, J. (April, 1994). *Classification error and goodness-of-fit in IRT models*. Paper presented at the annual meeting of the American Educational Research Association, April, New Orleans.

Stone, C.A. (October, 1993). *Assessing fit of the graded IRT model in the context of performance assessment*. Invited presentation at Educational Testing Service, Princeton.

Stone, C.A. (July, 1993). *The use of multiple replications in IRT based Monte Carlo research*. Paper presented at the European Meeting of the Psychometric Society, Barcelona, Spain.

Lane, S., Liu, M., Stone, C.A., & Ankenmann, R.D. (April, 1993). *Validity evidence for QUASAR's mathematics performance assessment.* Paper presented at the annual meeting of the American Educational Research Association, April, Atlanta.

Stone, C.A., Ankenmann, R.D., Lane, S., & Liu, M. (April, 1993). *Scaling QUASAR's performance assessment*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta.

Lane, S. & Stone, C.A. (April, 1992). *Empirical evidence for the reliability and validity of performance assessments*. Paper presented at the annual meeting of the National Council for Measurement in Education, San Francisco.

Ankenman, R.D. & Stone, C.A. (April, 1992). *A Monte Carlo study of marginal maximum likelihood estimates for the graded model*. Paper presented at the annual meeting of the National Council for Measurement in Education, San Francisco.

Stone, C.A. (April, 1991). *IRT based Monte Carlo research: How many replications*? Paper presented at the annual meeting of the National Council for Measurement in Education, Chicago.

Stone, C.A. (November, 1990). *Statistical computing packages for microcomputers*. Interview on "Computer Affairs", Pittsburgh Cable TV. Aired November.

Stone, C.A. (April, 1990). *An evaluation of marginal maximum likelihood estimates via the EM algorithm in the 2-parameter logistic response model*. Paper presented at the annual meeting of the National Council for Measurement in Education, Boston.

Lane, S., Stone, C.A., & Hsu, T. (April, 1990). *Diagnosing students' errors in solving algebra word problems*. Paper presented at the annual meeting of the National Council for Measurement in Education, Boston.

Stone, C.A. (April, 1987). *On the robustness of total indirect effects in the Jöreskog-Keesling-Wiley covariance structure model*. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.

Bergan, J.R., & Stone, C.A. (April, 1987). *Item parameter stability in IRT models for hierarchical skill sequences*. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.

Stone, C.A. (June, 1985). *LISREL VI*. Workshop presented to faculty and staff of the College of Nursing, University of Arizona.

Bergan, J.R., & Stone, C.A.  (April, 1985).  *Restricted item response models for developmental assessment*.  Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois.

Stone, C.A.  (April, 1985*).  A computer program for calculating total indirect effects and their standard errors in the Jöreskog covariance structure model*.  Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois.

Johnson, D.M., Parra, E.B., Anderson, D.O, & Stone, C.A.  (April, 1985).  *A study of item bias on six bilingual measures of preschool development*.  Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois.

Bergan, J.R., Feld, J.K., & Stone, C.A. (August, 1983). *Consultation and instruc­tional planning*. Paper presented at a symposium at the annual meeting of the American Psychological Association, Los Angeles, California.

Bergan, J.R., Stone, C.A., & Feld, J.K.  (August, 1982).  *Rule replacement in the development of basic number skills*. Paper presented at a symposium at the annual meeting of the American Psychological Association, Washington.

Stone, C.A., & Miller, S.L.  (June, 1980).  *A comparative analysis of attitudes towards male and female social roles in male and female high school students*.  Paper presented at the Rocky Mountain Psychological Association Convention, Tucson, Arizona.

Stone, C.A., & Miller, S.L.  (June, 1980).  *A comparative analysis of attitudes toward male and female social roles in male and female junior high students*.  Paper presented at the Rocky Mountain Psychological Association Convention, Tucson, Arizona.

**TECHNICAL REPORTS**

Lane, S., Stone, C., Draper, J., & Krost, K. (June 2013). *Pennsylvania Educator Effectiveness Project 2012-201: Results from the Intermediate Unit Leads.* Pittsburgh, PA: University of Pittsburgh. (Award # 4300345494).

Lane, S., Stone, C., Draper, J., Scott, P. & Krost, K. (January 2013). *Pennsylvania Teacher Evaluation 2011-2012 Pilot: Report for the Teachers and Principals/Evaluators.* Pittsburgh, PA: University of Pittsburgh. (Award # 4300345494).

Lane, S., Stone, C., & Blaze, T. (September 2012). *Pennsylvania Teacher Evaluation 2011-2012 Pilot: Report for the IU Leads and Trainers.* Pittsburgh, PA: University of Pittsburgh. (Award # 4300345494).

Stone, C.A., & Tierney, S. (2009). Student Achievement Report Year Six: Math and Science PSSA Scores 2003-2008. In Tanannis, C. (Eds.), *Evaluator’s Report Year Six: Math and Science Partnership of Southwestern Pennsylvania*. Report prepared for the National Science Foundation (Grant Number EH4 0314914).

Stone, C.A., Ye, F., Lane, S., Zhu, X.I., Scholl, L.H. (November, 2006). *Providing subscale scores for diagnostic information: An evaluation of different methods for augmenting scores*. Final report prepared for the Delaware State Department of Education.

Stone, C.A. & Zhu, X. (September, 2006). *Evaluating the ACTS of PA curriculum intervention: A meat-analysis using a quasi-experimental design*. Final Report prepared for SERVE, University of North Carolina at Greensboro.

Stone, C.A. & Zhu, X. (March, 2006). *Evaluating the ACTS of PA curriculum intervention: Implementing propensity scoring in a quasi-experimental design*. Final Report prepared for SERVE, University of North Carolina at Greensboro.

Stone, C.A. & Weisman, A. (August, 2002). *Simulating computer adaptive tests*. Final Report for University of Pittsburgh, Central Research Development Fund.

Stone, C.A. & Dresher, A. (August, 2001). *Evaluation of a Wald test for assessing the significance of a fit statistic based on posterior expectations in item response theory models*. Final report for the Department of Education, NAEP Secondary Analysis Program (PR/Award Number R902B000008).

Stone, C.A. and Hansen, M.A. (May, 2000). *A computer program for assessing goodness-of-fit of item response theory models to NAEP data*. Final report for the Department of Education, NAEP Secondary Analysis Program (PR/Award Number R902B70008).

Lane, S., Parke, C.S., Stone, C.A., & Cerrillo, T.L. (Dec, 2000). *MSPAP impact study: Science*. Final Report prepared for the U.S. Department of Education, Assessment Development and Evaluation Grants Program (CFDA 84.279-A), for the Maryland Assessment System Project.

Lane, S., Parke, C.S., Stone, C.A., & Hansen, M.A. (Nov, 2000). *MSPAP impact study: Social Studies*. Final Report prepared for the U.S. Department of Education, Assessment Development and Evaluation Grants Program (CFDA 84.279-A), for the Maryland Assessment System Project.

Lane, S., Parke, C.S., Stone, C.A., & Cerrillo, T.L. (Sept, 2000). *MSPAP impact study: Language Arts*. Final Report prepared for the U.S. Department of Education, Assessment Development and Evaluation Grants Program (CFDA 84.279-A), for the Maryland Assessment System Project.

Lane, S., Parke, C.S., & Stone, C.A. (Dec, 1999). *MSPAP impact study: Mathematics*. Prepared for the U.S. Department of Education, Assessment Development and Evaluation Grants Program (CFDA 84.279-A), for the Maryland Assessment System Project.

Stone, C.A. & Lane, S. (1987). *CAP Manual for Test and Item Analysis*. Prepared for DeVRY Inc., Illinois.

Bergan, J.R., Lane, S., Anderson, D.O., Swarner, J.C., & Stone, C.A. (November, 1985). *Head Start Measures Project Program Progress Report for 1984 - 1985*. Prepared for the Department of Health and Human Services Contract No. HHS‑90‑CD‑0524/01.

Bergan, J.R., Anderson, D.O., Feld, J.K., Henderson, R.W., Johnson, D.M., Lane, S., Mott, S.E., Parra, E.B., Robinson, L.A., Stone C.A., & Swarner, J.C. (October, 1984).  *Head Start Measures Project Final Report*. Prepared for the Department of Health and Human Services Contract No. HHS‑105‑81‑C‑008.

Bergan. J.R., Stone, C.A., Anderson, D.O., Feld, J.K., Johnson, D.M., Lane, S., Parra, E.B., & Swarner, J.C.  (March, 1984).  *Head Start Measures Technical Manual*. Prepared for the Department of Health and Human Services, Contract No.:  HHS‑105‑81‑C‑008.

Bergan, J.R., Feld, J.K., Johnson, D.M., Henderson, R.W., Mott, S.E., Stone, C.A., Anderson, D.O., Swarner, J.C., Schnaps, A., Lane, S., & Rosser, R. A.  (March, 1984).  *Head Start Measures Battery Examiner`s Manual.* Prepared for the Department of Health and Human Services, Contract No.:  HHS‑105‑81‑C‑008.

Bergan, J.R., Stone, C.A., Feld, J.K., Schnaps, A., & Anderson, T.A. (1982). *Math dimension measure:  Item tryout version for the Head Start Program*. Prepared for the Department of Health and Human Services, Contract No. 105‑81‑C‑008.

Bergan, J.R., Stone, C.A., Henderson, R.W., Rosser, R.A., Silvers, P.J., Hestand, J., & Feld, J.K.  (1982).  *The development of path‑referenced cognitive measures of Head Start:  Item tryout report*.  Prepared for the Department of Health and Human Services, Contract No. 105‑81‑C‑008.

Bergan, J.R., Stone, C.A., Cancelli, A.A., Anderson, T.A., & Schnaps, A. (1981).  The developmental structure of early mathematical knowledge: A measurement perspective. Prepared for the Department of Health and Human Services, Contract No. 105‑81‑C‑008.

Bergan, J.R., Cancelli, A.A., Stone, C.A., & Towstopiat, O.  (1980).  *A structural approach to the validating of hierarchical training sequences: A final report*. Prepared for the Defense Advanced Research Projects Agency, Contract No. MDA903‑70‑C‑0912.