DOCUMENT RESUME

ED 399 287 TM 025 515

AUTHOR Dimitrov, Dimiter M.
TITLE Monte Carlo Approach for Reliability Estimations in Generalizability Studies.
PUB DATE Apr 96
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classification; Criterion Referenced Tests; Cutting Scores; *Estimation (Mathematics); *Generalizability Theory; *Monte Carlo Methods; Probability; *Scores; *Test Reliability
IDENTIFIERS *Probabilistic Models; Statistical Analysis System

ABSTRACT

A Monte Carlo approach is proposed, using the Statistical Analysis System (SAS) programming language, for estimating reliability coefficients in generalizability theory studies. Test scores are generated by a probabilistic model that considers the probability for a person with a given ability score to answer an item with a given difficulty parameter correctly. Three types of reliability-like coefficients available in generalizability theory are considered: (1) the generalizability coefficient that is the analog of the classical reliability coefficient; (2) the index of dependability, appropriate for criterion-referenced testing; and (3) the classification reliability index, for the reliability of classification decisions based on a cutting score. The proposed approach is illustrated for a single-facet crossed design but it works for higher levels of crossed or nested generalizability designs. The SAS program allows flexibility and control on factors such as the type of probabilistic model, the type of ability score and difficulty parameter distributions, the location of the cutting score, and the amount of information provided by each item. An appendix presents the SAS program for Monte Carlo reliability estimations. (Contains one table and eight references.) (SLD)

ED 399 287

MONTE CARLO APPROACH FOR RELIABILITY ESTIMATIONS IN GENERALIZABILITY STUDIES

Dimiter M. Dimitrov

Kent State University

A paper presented at the Annual Meeting of
The American Educational Research Association
New York, April 8-12, 1996

MONTE CARLO APPROACH FOR RELIABILITY ESTIMATIONS IN GENERALIZABILITY STUDIES

OBJECTIVES

The purpose of the present paper is to propose a Monte Carlo approach, by using the SAS programming language, for estimation of reliability coefficients in generalizability theory studies. Test scores are generated by a probabilistic model of the type $P_{ip} = F(\theta_p, b_i)$, where P_{ip} is the probability for person **p** with ability score θ_p to answer correctly item **i** with difficulty parameter b_i. Under consideration are three types of reliabilty-like coefficients available in generalizability theory: (a) Generalizability coefficient, G_p, which is the analog of the classical reliability coefficient, (b) Index of dependability, Φ, which is appropriate for criterion-referenced testing, and (c) Classification reliability index, $\Phi(\lambda)$, for reliability of classification decisions based on a cutting score, λ, (Brennan and Kane, 1977). The Monte Carlo approach is illustrated here for a single-facet crossed design but it works for higher levels of crossed and/or nested generalizability designs. The way it is designed, the SAS program allows flexibility and control on factors such as: (a) type of the probabilistic model $F(\theta_p, b_i)$, (b) type of the θ_p and b_i distributions, (c) location of the cutting score, λ, and (d) the amount of information provided by each item, $I_i(\theta)$.

FRAMEWORK

In the classical approach of measurement, under the attempt to standardize measurement condi- tions, there are only two sources of score variation - variance across people (true variance) and variance of unspecified sources (error variance). In the generalizability theory approach of mea- surement it is possible to asses multiple sources of variation. For example, in addition to items, raters may be involved as a measurement condition which is permitted to vary. The variance across the object of measurement is called **true variance** and the variance associated with all other sources is called **error variance. The** sources of measurement error are referred to as **facets** of measurement. The process through which measurement scores are gathered may be crossed or nested. In a Generalizability theory study **(G-study),** the purpose is to estimate the variances associated with various facets of measurement. Further, consequences of various changes in the measurement design can be investigated to seek the optimal design. For example, one may investigate the reduction of error as a result of treating some of the facets as fixed. The investigations of various designs are referred to as Decision studies **(D-studies).** In the classical test theory all facets are generally assumed standardized and the only facet is items. This is analogous to a G-study in which the concern is with the variance of the total score.

The major statistical difference between a crossed design and a nested design is that the main effect variance component due to the nested facet is not separable statistically from the interaction variance component through ANOVA. This fact leads to more conservative generalizability coefficients in the **D-studies** than their crossed design counterparts. For an **i(p)** design, items **(i)** nested within persons **(p)**, the variance of all observed scores, X_{ip}, is given by:

$$\sigma^2(X_{ip}) = \sigma_p^2 + \sigma_{i(p)}^2 \tag{1}$$

For a single-facet crossed design **p x i**, where persons are crossed with items, the variance of the observed scores, X_{ip}, is given by:

$$\sigma^2(X_{ip}) = \sigma_p^2 + \sigma_i^2 + \sigma_{ip}^2 \tag{2}$$

Related to the reliability coefficients under consideration in this paper, G_ρ, Φ, and $\Phi(\lambda)$, are also the concepts of **absolute error variance**, $\sigma^2(\Delta)$, and **relative error variance**, $\sigma^2(\delta)$. The **absolute error** is the difference between a person's observed and universe scores, $\Delta_p = X_{ip} - \mu_p$, and the **relative error** is the difference between a person's observed deviation score and his/her universe deviation score, $\delta_p = (X_{ip} - \mu_i) - (\mu_p - \mu)$.

If $\sigma^2(\tau)$ stands for the variance of the object of measurement, τ, in a G-study, the reliability coefficients G_ρ, Φ are defined as follows:

$$G_\rho = \frac{\sigma^2(\tau)}{\sigma^2(\tau)+\sigma^2(\delta)} \tag{3}$$

$$\Phi = \frac{\sigma^2(\tau)}{\sigma^2(\tau)+\sigma^2(\Delta)} \tag{4}$$

If person's achievement is the only object of measurement, $\tau = p$, then $\sigma^2(\tau) = \sigma_p^2$ and the classification reliability index is defined by:

$$\Phi(\lambda) = \frac{\sigma_p^2+(\mu-\lambda)^2}{\sigma_p^2 + \sigma^2(\Delta)+(\mu-\lambda)^2} \tag{5}$$

An unbiased estimate of $(\mu - \lambda)^2$ is $(\overline{X} - \lambda)^2 - \sigma^2(\overline{X})$, where $\overline{X}$ is the mean test score of n_p persons on n_i test items.

For example, for one-facet crossed design "persons x items" the mean score variance is given by:

$$\sigma^2(\overline{X}) = \sigma_p^2/n_p + \sigma_i^2/n_i + \sigma_{ip}^2/(n_p n_i) \tag{6}$$

and the absolute error variance is given by:

$$\sigma^2(\Delta) = \sigma_i^2/n_i + \sigma_{ip}^2/n_i \tag{7}$$

By taking into account (6) and (7), the following unbiased estimate of the classification reliability index can be obtained from (5) for one-facet crossed design:

$$\hat{\Phi}(\lambda) = \frac{\hat{\sigma}_p^2 + (\overline{X}-\lambda)^2 - \hat{\sigma}^2(\overline{X})}{\hat{\sigma}_p^2 + \hat{\sigma}^2(\Delta) + (\overline{X}-\lambda)^2 - \hat{\sigma}^2(\overline{X})} \tag{8}$$

METHOD

Developed was a relatively simple **SAS** program for Monte Carlo simulations and calculation of the reliability indices G_ρ, Φ, and $\Phi(\lambda)$ from formulas (3), (4) and (5), respectively. It assumes an error-free measure of person's ability, θ_p, on the same scale of measurement for item's difficulty, b_i, and allows different simulations of the θ_p and b_i distributions. It also allows different simulations of the probabilistic model $P_{ip} = F(\theta_p, b_i)$ for calculating the probability that a person with ability score θ_p will answer correctly item with difficulty b_i. The user of the SAS program has the freedom to set different cut-off scores, λ, for the calculation of $\Phi(\lambda)$, and to select simulated items that provide information above some specified amount, $I_i(\theta)$.

The results, reported later in this paper, were obtained from Monte Carlo simulations of different IRT models. For the purposes of unification and comparisons, same number of replications (NR = 20), same number of persons (NP = 100), and same number of items (NI = 20) were used for each IRT model. Hence, 40000 observations were manipulated with each execution of the respective SAS program. Used were the following illustration models:

(A) Two-parameter IRT model: Both person's ability, θ_p, and item difficulty, b_i, were taken from the standard normal distribution, by using the SAS function RANNOR, and substituted in the one-parameter IRT model:

$$P_{ip} = \{1 + \exp[(D)(a_i)(\theta_p - b_i)]\}^{-1} \tag{9}$$

the item discrimination parameter, $\mathbf{a}_i$, is proportional to the slope of the item characteristic curve (**ICC**) at the point $\mathbf{b}_i$ on the ability scale. The constant **D** was fixed to 1.7 which made the model in (9) almost the same as the normal-ogive model (see, e.g., Hambleton et al., 1991, p. 15). The score, X_{ip}, of person **p** on item **i** was simulated by the **SAS** function RANBIN(0, 1, P_{ip}). The **ANOVA** procedure was used for the statistics involved in the single-facet crossed design.

(B) One-parameter IRT model: The simulations of this model were conducted under the same conditions as for the two-parameter model in (9), with the difference that the item discrimination parameter was the same for all items ($a_i = a$):

$$P_{ip} = \{1 + \exp[1.7(a)(\theta_p - b_i]\}^{-1} \tag{10}$$

The results were compared for three different values of the fixed parameter, a = .1, a = .5, and a = 1.8, looking for possible changes in the reliability coefficients G_p, Φ, and $\Phi(\lambda)$.

(C) Rasch's Poisson model: The general form of this model (see, e.g., Allen & Yen, 1979, p.250), was reduced here to the calculation of probability of person j making no error on (i.e. giving correct answer to) an item i:

$$P(X_{ij} = 0 | \lambda_{ij}) = \exp(-\lambda_{ij}), \tag{11}$$

where $\lambda_{ij} = \frac{b_i}{\theta_j}$. The item difficulty, b_i, and person's ability, θ_j, were generated randomly from the uniform distribution on the interval (0,1), by using the SAS function UNIFORM.

(D) Chi-square model: This model was used for the illustration of a hypothetical measurement situation where, for example, the person's ability level is fixed and the probability for a given dichotomous response, P_{ij}, decreases when the product of factors such as person's anxiety level, ξ_j , and task difficulty, τ_i , increases. Then, the respective function $P_{ij} = F(\theta_j, \xi_j, \tau_i)$ may be approximated, say, by the chi-squared probability distribution function. This hypothetical situation is illustrated in the current study with the simulation of P_{ij} by the use of the SAS function PROBCHI for chi-squared probability distribution, with degrees of freedom fixed to df = 3.

(E) Uniform model: With this model, the probability P_{ij} for person j to answer correctly item i was generated as a pseudo-random variate uniformly distributed on the interval (0,1), by using the SAS function UNIFORM.

After determining the P_{ij} value in the any of the above models, a dichotomous response of 0 or 1 was generated as an observation from the binomial distribution B(n, x, p), with n = 1, x = 1, and p = P_{ij} , by using the SAS function RANBIN. The entire Monte Carlo procedure, with 20 replication for 100 persons and 20 items, was executed for 5 different cut-off score values, λ , for the analysis of the classification reliability, $\Phi(\lambda)$; (λ = 6, 8, 10, 12, 14). These cut-off scores are symmetrically located around the theoretical average score, μ = 10, on the 20 item test for the uniform and logistic models used in this study. For the illustrative examples, the distribution mean of the scores generated from the chi-square probability model was close to μ = 12, and the score mean for the Rasch's Poisson model was close to μ = 8.

RESULTS

Table 1 summarizes the results from the Monte Carlo simulation of the six models described above in (A) - (E). A general inspection of the reliability indices reveals consistency of the results with the respective theoretical relations between them. For instance, the generalizability coefficients, G_ρ, are systematically higher than the dependability indices, Φ. This can be seen from the comparison of formulas (3) and (4), taking into account that $\sigma^2(\Delta) > \sigma^2(\delta)$. Also, The comparison of classification indices, $\Phi(\lambda)$, calculated for different cut-off scores, λ , confirms the theoretical fact that the closer λ to the distribution mean, μ, the lower $\Phi(\lambda)$ - see formula (5). The standard deviations of the reliability indices, across their values from the 20 replications, were very small and varied between 0.02 and 0.0007. All those facts, along with the high values of all reliability indices, support the validation of the proposed Monte Carlo procedure.

From another perspective, the comparison of G_ρ and Φ indicates that the generalizability coefficient, G_ρ , is more resistant to the type of probability model. As it was mentioned at the beginning, G_ρ is the analog of the classical reliability coefficient, whereas Φ is appropriate for criterion-referenced testing.

In terms of $\Phi(\lambda)$, the Chi-square and the Uniform models are more sensitive to the position of the cut-off score, λ, compared to the other three models. Also, the one parameter logistic model leads to higher $\Phi(\lambda)$ values, compared to the two parameter logistic model, especially when λ is close to the population mean score (in this example, μ = 10 for the two logistic models).

CONCLUSION

The Monte Carlo approach developed in this study allows the calculation of all reliability indices, G_ρ, Φ, and $\Phi(\lambda)$, in G-studies, for different distributions of person's ability scores, θ_p, different distributions of item difficulty, b_i, and different IRT-type models for the probability of correct answer, $P_{ij} = F(\theta_p, b_i)$. One can also incorporate models for the probability of responses based on personality and/or task parameters other than θ_p and b_i, respectively. For the purpose of classification reliability estimations, one can vary the position of the cut-off score, λ, for different optimizations of the $\Phi(\lambda)$ index.

Technically, it is easy to set restrictive conditions about the amount of information provided by the simulated items, range of person's ability score, and other factors influencing the quality of the measurement. For example (see, e.g., Hambleton et al., 1991, p. 97), for the two parameter logistic model in (9), the amount of item information, $I(\theta_j)$, is calculated by a formula that is easy to incorporate in the SAS program given in the appendix:

$$I(\theta_j) = 2.89(a_i)(P_{ij})(1 - P_{ij})$$

In addition to the applications discussed above, one can use the variance components calculated by the SAS procedure (see Appendix) for other important purposes such as error reduction, confidence interval calculations, and standard settings in various D-study designs.

Table 1

Monte Carlo averaged estimates of GT reliability indices from 20 replications of simulated observations on 100 persons and 20 items

Probability model	Classification index, $\Phi(\lambda)$					Dependability index, Φ	Generalizability coefficient, G_ρ
	$\lambda = 6$	$\lambda = 8$	$\lambda = 10$	$\lambda = 12$	$\lambda = 14$		
2P logistic	.973	.927	.826	.926	.973	.848	.979
1P logistic	.975	.949	.924	.949	.976	.927	.982
Rasch' Poisson	.958	.929	.961	.983	.991	.932	.982
Chi-square	.984	.968	.912	.634	.786	.682	.975
Uniform	.973	.919	.740	.911	.972	.787	.976

Note. λ = cut-off score for classification decisions ; 2P logistic = Two parameter logistic model; 1P logistic = One parameter logistic model.

References

Allen, M. J. & Yen, W. M. (1979). Introduction to measurement theory. Monterey, Brooks/Cole Publishing Company.

Brennan, R.L., & Kane, M.T. (1977). An index of dependability for mastery test. Journal of educational measurement, 14, 271-289.

Cronbach, L. J., Nanda, G.C., & Rajarathnam, N. (1972). The dependability of behavioral measurements:Theory of generalizability for scores and profiles. N.Y.: Wiley.

Feldt, L.S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), Educational measurement (3rd ed.), pp. 105-146. New York: Macmillan.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory. Newbury Park SAGE Publications.

Livingston, S.A.(1972). Criterion-referenced applications of classical test theory. Journal of Educational Measurement, 9, 13-26.

Shavelson, R.J.,& Webb, N. M. (1991). Generalizability theory. A Primer. Newbury Park: SAGE Publications.

Smith, J. K. & Cody, R. P. (1991). Applied statistics and the SAS programming language. N.Y., North-Holland.

Suen, H. K. (1990). Principles of test theories. N.J.: Lawrence Erlbaum

APPENDIX

SAS PROGRAM FOR MONTE CARLO RELIABILITY ESTIMATIONS IN G-STUDIES
(An illustration version for the two-parameter logistic model)

```
==========================================================================
DATA CROSSED;
NR = 20;
NP = 100;
NI = 20;
NRL1 = 3*NR - 2;
NRL2 = NRL1 + 1;
NRL3 = NRL2 +1;
DO REP = 1 TO NR;
DO PERS = 1 TO NP;
TP=RANNOR(0);
scrp=0;
DO ITEM= 1 TO NI;
TI= RANNOR(0);
A = UNIFORM(0);
E = - 1.7*(TP-TI)*A;
P = 1/(1 + 2.7182818**E);

SCOR = RANBIN(0,1,P);
scrp = scrp + scor;
OUTPUT;
END;
END;
END;
PROC SORT; BY REP;
PROC ANOVA OUTSTAT= POST;
CLASSES PERS ITEM; BY REP;
MODEL SCRP= PERS ITEM PERS*ITEM/nouni;
OUTPUT OUT= POST SS DF;
RUN;

data numb;
set crossed(KEEP=NR NP NI NRL1 NRL2 NRL3);
run;

data score;
set crossed;
scorepi = scrp;
run;

data adjscore;
set numb;
NPR = NR*NP*NI;
do n = NI to NPR by NI;
set score point=n;
if _error_=1 then abort;
output;
end;
stop;
```

APPENDIX (cont.)

==

```
DATA KUKA;
SET adjscore;
PROC SORT;BY REP;
PROC MEANS;
 VAR SCOREPI; BY REP;
 OUTPUT OUT=BARR MEAN=XBAR;
RUN;

DATA TWO;
SET POST;
IF DF > 0 THEN MS=SS/DF; ELSE DELETE;
RUN;

DATA TWO1;
SET TWO;
MSP=MS;
RUN;

DATA TWO2;
SET TWO;
MSI=MS;
RUN;

DATA TWO3;
SET TWO;
MSPI=MS;
RUN;

DATA PERS;
set numb;
DO N=1 TO NRL1 BY 3;
 SET TWO1 POINT=N;
 IF _ERROR_=1 THEN ABORT;
 OUTPUT;
END;
STOP;
RUN;

DATA ITEM;
set numb;
DO N = 2 TO NRL2 BY 3;
  SET TWO2 POINT = N;
  IF _ERROR_=1 THEN ABORT;
  OUTPUT;
END;
STOP;
RUN;
```

APPENDIX (cont.)

```
==========================================================================

DATA PERSITEM;
set numb;
DO N = 3 TO NRL3 BY 3;
   SET TWO3 POINT=N;
   IF _ERROR_=1 THEN ABORT;
   OUTPUT;
END;
STOP;
RUN;

DATA RESUL;
set numb;
MERGE PERS ITEM PERSITEM ;
NPI=NP*NI;
SGMP2=(MSP-MSPI)/NI;
IF SGMP2 < 0 THEN SGMP2=0;
SGMI2=(MSI-MSPI)/NP;
IF SGMI2<0 THEN SGMI2=0;
SGMPI2=MSPI;
SX2=SGMP2/NP + SGMI2/NI + SGMPI2/NPI;
GR=SGMP2/(SGMP2 + SGMPI2/NI);
SGMDLT2 = SGMI2/NI + SGMPI2/NI;
FI=SGMP2/(SGMP2 + SGMDLT2);
RUN;

DATA KP;
SET RESUL(KEEP=FI GR SX2 SGMDLT2 SGMP2);
RUN;

DATA FNL;
MERGE BARR KP;
DO LBD = 4 TO 14 BY 2;
DEV2 = (XBAR - LBD)*(XBAR -LBD) - SX2;
FIL = (SGMP2 + DEV2)/(SGMP2 + SGMDLT2 + DEV2);
OUTPUT;
END;
proc sort; by LBD;
proc means; var xbar gr fi fil; by lbd;
RUN;
========================================================================
```

Note: All changes related to the use of different probability models, different simulations of person's ability and/or item difficulty, and different restrictive conditions on the model, concern only the first data set (DATA CROSSED). They should be done by using the respective SAS manipulations, whithout changing the other data sets. Changes related to the cut-off score values concern only the DO command in the last data set (DATA FNL).

TM025575

AERA April 8-12, 1996

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: MONTE CARLO APPROACH FOR RELIABILITY ESTIMATIONS IN GENERALIZABILITY STUDIES

Author(s): DIMITER M. DIMITROV

Corporate Source:

Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

☑ Sample sticker to be affixed to document	Sample sticker to be affixed to document ☐
Check here Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction	**or here** Permitting reproduction in other than paper copy.
"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."	"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."
Level 1	**Level 2**

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: [signature]	Position: Assistant Professor
Printed Name: DIMITER M. DIMITROV	Organization: KENT STATE UNIVERSITY
Address: 507 WHITE HALL, COLLEGE OF EDUCATION, KENT STATE UNIVERSITY, KENT, OHIO 44242-0001	Telephone Number: (330) 672-7918
	Date: 04/09/96

CUA

THE CATHOLIC UNIVERSITY OF AMERICA
Department of Education, O'Boyle Hall
Washington, DC 20064
202 319-5120

February 27, 1996

Dear AERA Presenter,

Congratulations on being a presenter at AERA[1]. The ERIC Clearinghouse on Assessment and Evaluation invites you to contribute to the ERIC database by providing us with a written copy of your presentation.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality.

Please sign the Reproduction Release Form on the back of this letter and include it with **two** copies of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can drop off the copies of your paper and Reproduction Release Form at the **ERIC booth (23)** or mail to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to:

AERA 1996/ERIC Acquisitions
The Catholic University of America
O'Boyle Hall, Room 210
Washington, DC 20064

This year ERIC/AE is making a **Searchable Conference Program** available on the AERA web page (http://tikkun.ed.asu.edu/aera/). Check it out!

Sincerely,

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

---

[1]If you are an AERA chair or discussant, please save this form for future use.