DATA SHARE AND USE AGREEMENT
between
ILLINOIS STATE BOARD OF EDUCATION
and
THE NATIONAL OPINION RESEARCH CENTER

This Data Share and Use Agreement (hereinafter "Agreement") is hereby made and entered into by and between the Illinois State Board of Education ("ISBE") and the National Opinion Research Center ("NORC") (the "Recipient") (each a "Party" and collectively "the Parties").

I. RECITALS

1. ISBE owns and maintains information, including individually identifiable information, on students, including prior Illinois students ("Student Data"); teachers, including teacher certification and service record data ("Teacher Data"); programs; schools or institutions; and districts (collectively "Confidential Data") necessary for required federal reporting and to audit and evaluate education programs and to perform studies for, or on behalf of, public elementary and secondary schools, all in a manner consistent with the Family Educational Rights and Privacy Act (20 U.S.C. § 1232g) ("FERPA"), the Illinois School Students Records Act (105 ILCS 10/1, *et seq.*) ("ISSRA"), and other applicable laws.

2. ISBE owns and maintains the Student Information System ("SIS"), a state-level student information system which assigns all students throughout the State a unique identification number which allows for the capacity to follow a student's progress over time and collects data, including data on the student's enrollment, attendance, and progress.

3. The Data are and at all times will remain the sole property of ISBE. ISBE retains all right, title and interest in and to the Data and all copies thereof (including, without limitation, all copyrights, trade secrets, trademarks, patents, and other similar proprietary rights therein).

4. The term "individually identifiable information" means information that is identifiable to a particular individual, program, classroom, school, institution or district, including but not limited to the following: (a) a first and last name; (b) a home or other physical address, including street name and name of a city, town, or county; (c) an e-mail address; (d) a telephone number; (e) a social security, employer identification, or student identification number; (f) test scores; or (g) clinical information, including any questionnaires, notes, or other documentation.

5. ISBE is authorized by law to secure, compile, catalog, publish and preserve information and data relative to the public school system of Illinois, making such comparison as will assist the General Assembly in determining the priorities of educational programs to be of value to the public school system of Illinois and of other states (105 ILCS 5/2-3.31).

6. ISBE is the state agency in Illinois responsible for educational policies and guidelines for public schools, pre-school through grade 12, and Vocational Education in Illinois and is responsible for analyzing the present and future aims, needs and requirements of education in Illinois (105 ILCS 5/1A-4 (C)).

7. Recipient is NORC, an independent 501 (c) 3 research organization headquartered in downtown Chicago with additional offices on the University of Chicago's campus, the D.C. Metro area, Atlanta, Boston, and San Francisco, that serves the public interest through objective social science research that supports informed decision-making in key areas such as education, health, economics, justice and the environment. NORC is entering into this Agreement on behalf of its Director and Senior Research Scientist, Kathleen Parks, and employee, Eric Hedberg, NORC Senior Research Scientist.

8. Recipient will engage in research and evaluation of the Confidential Data as ISBE's authorized representative. This research and evaluation, as more fully set forth in Exhibit A, the Specifications for Shared Data and Use ("Specifications"), attached hereto, will evaluate Advancing State-specific Design Parameters for Designing Better Evaluation Studies, as defined in Section A. of the Specifications,("Research Project").

9. Federal law allows the release of educational records or personally identifiable information of students without the consent of students or parents ("individually identifiable student information"), so long as the disclosure is to authorized representatives of Federal or state educational authorities for purposes of audit or evaluation of Federal or state supported education programs, or for the enforcement of or compliance with Federal legal requirements that relate to those programs. or to organizations or individuals conducting studies for, or on behalf of, educational agencies or institutions for the purpose of improving instruction or developing, validating or administering predictive tests only if—

   (A) The study is conducted in a manner that does not permit personal identification of parents and students by individuals other than representatives of the organization that have legitimate interests in the information;
   (B) The information is destroyed when no longer needed for the purposes for which the study was conducted; and
   (C) The educational agency or institution or the State or local educational authority enters into a written agreement with the organization. 20 U.S.C. § 1232g(b)(1)(C)(i)(III) and (b)(1)(F) and 34 C.F.R. § 99.31 (a)(3)(iv) and (a)(6)(i),(a)(6)(ii), (a)(6)(iii) and § 99.35.

10. The Parties wish to enter into this Agreement in order to:

    A. Establish Recipient as ISBE's authorized representative for purposes of Recipient providing ISBE with research, analysis, audit and/or evaluation of the State's educational system for the improvement of educational instruction;

    B. Establish the necessary data sharing arrangements between the Parties to provide data necessary to conduct research, analysis, and evaluation;

C. Facilitate the audit or evaluation of education programs for, or on behalf of, ISBE in a manner permitted by FERPA, ISSRA, and other applicable law ("Audit or Evaluation"); and

D. Facilitate the performance of studies for, or on behalf of, ISBE in a manner permitted by FERPA, ISSRA, and other applicable law ("Research").

11. ISBE will, in light of the December 2011 guidance from the U.S. Department of Education, post for public access, a copy of this Agreement within the Data Request History located on its website.

II. DATA SHARING AUTHORIZATIONS

1. ISBE hereby designates and recognizes Recipient as its authorized representative for purposes of Research, Audit and Evaluation related to Advancing State Specific Design Parameters forDesigning Better Evaluation Studies. Illinois has a rich community of scholars and education researchers who perform experiments or other evaluations to test the curriculum that instructs Illinois students. However, the value of these studies depends on the ability to draw conclusive evidence from the samples used. Without the estimates provided by this study, it is difficult to ascertain whether the study designs can give conclusive evidence. Thus, in order to prevent wasting time and resources, and to help future research be more cost effective and valid, it is vital to the Illinois research community that these parameters are estimated and published. For a detailed description of the project, see Exhibit A.

2. ISBE agrees to share with Recipient the ISBE data (hereafter referred to as "Confidential Data") set forth in the Specifications solely for the limited purposes and extent as specified in Exhibit A. Recipient agrees to use the Confidential Data for the Research Project only as authorized pursuant to this Agreement. Any act by Recipient that involves a use beyond that set forth herein shall be deemed in its entirety to be a prohibited use of the Confidential Data.

3. Confidential Data include any temporary data analysis files or subsets of the original data file that contains any personally identifiable information.

4. Recipient will securely store and maintain the Confidential Data at NORC secure servers at North State St., 16th Floor, Chicago, IL 60602, in accordance with the terms and conditions set forth in Exhibit E ("Data Security Plan"), attached hereto.

5. All officials and employees authorized to request, receive and obtain information, including the Confidential Data, from Recipient under this Agreement are identified in Exhibit A, Attachment 1. Any further disclosure to officials and employees identified by Recipient that will have access to information as provided under this Agreement shall be documented through completion of a form meeting the requirements of Exhibit C ("Form for Amendment of Exhibit A, Attachment 1") and an executed copy(ies) of Exhibit E, Attachment 1 ("Security Pledge for the Use of Confidential Data"). No disclosure of information provided under this

Agreement shall be permitted until a form meeting the requirements of Exhibit C and an executed Security Pledge for the Use of Confidential Data is received and approved by ISBE.

6. All contractors, subcontractors, or agents authorized to request, receive or obtain information, including Confidential Data, from Recipient under this Agreement are identified in Exhibit A, Attachment 2. Any further disclosure to contractors, subcontractors, and agents identified by Recipient that will have access to information as provided under this Agreement shall be documented through completion of a form meeting the requirements of Exhibit D ("Form for Amendment of Exhibit A, Attachment 2") and an executed copy(ies) of Exhibit E, Attachment 1 ("Security Pledge for the Use of Confidential Data"). No disclosure of information provided under this Agreement shall be permitted until a form meeting the requirements of Exhibit D and an executed Security Pledge for the Use of Confidential Data is received and approved by ISBE.

7. Recipient agrees that it is ultimately responsible for ensuring that any third-party, including any employee, contractor, subcontractor, or agent of Recipient, operates in accordance with the terms and conditions of this Agreement.

8. Recipient is responsible for ensuring that any third party acquirer of ISBE Confidential Data, employed by, under contract to or working in collaboration with Recipient, operate within the terms of this Agreement; not retain any Confidential Data permanently, but is required to return any Confidential Data to ISBE upon request, or to Recipient within 10 days of Recipient's request, and/or upon completion of the work or termination, cancellation or expiration of this Agreement. ISBE may request that the Confidential Data be permanently destroyed by an approved method outlined in Exhibit E and that written certification of destruction be sent to ISBE.

9. ISBE makes no representations or warranties, express or implied, with respect to the Confidential Data. ISBE shall not be liable to the Recipient for amounts representing the loss of profits, loss of business or indirect, consequential or punitive damages in connection with the provision or use of Confidential Data under this Agreement.

10. The individuals signing below on behalf of ISBE represent that, with respect to any agreement between any third-party and the Recipient, ISBE's signatures attached hereto are intended solely as an acknowledgement of the separate data agreements with the third-party and do not suggest or imply acceptance of the terms and conditions of any agreements between the Recipient and the third-party, nor constitute an endorsement or approval of any such agreements by the State of Illinois.

III. DATA ACCESS, USE AND SECURITY

1. **Restrictions on Recipient.** The data access, use, and security restrictions set forth in this Section shall apply to the receipt, use, disclosure, and maintenance of Confidential Data by Recipient. Recipient agrees to the following:

A. Confidential Data may only be used for the purpose or purposes authorized pursuant to this Agreement.

B. Recipient will comply with all applicable laws, materials, regulations and all other State and Federal requirements with respect to the protection of privacy, security and dissemination of the shared data.

C. Recipient will comply with the relevant requirements of FERPA (20 U.S.C. § 1232g) and ISSRA (105 ILCS 10/1 et seq.), regarding the confidentiality of Student Data, and specifically "education records" as defined in FERPA and "school student records" as defined in ISSRA. Any use of information contained in student education records to be released must be approved by ISBE. To protect the confidentiality of student education records, Recipient will limit access to student education records to those employees who reasonably need access to them in order to perform their responsibilities under this Agreement.

D. Recipient will follow ISBE's confidentiality requirements for all ISBE data, pursuant to the Data Processing Confidentiality Act (30 ILCS 585/0.01 et seq.). Information obtained from any individual shall comply with the following terms and conditions, which include, but are not limited to:
   - Be confidential;
   - Not be published or open to public inspection;
   - Not be used directly in any court in any pending action or proceeding; and
   - Not be admissible in evidence in any action or proceeding.

   All records and other information maintained by ISBE regarding any person are confidential and shall be protected from unauthorized use and/or disclosure under this Agreement. Any dissemination or use of the Confidential Data for other than the primary purpose of this Agreement without the express written authority of ISBE is specifically prohibited. Confidential Data released under this Agreement are solely for the use of Recipient and are to be used only for the specific purposes as described in the Specifications.

E. In the event that any Confidential Data is required to be disclosed in response to a valid order of a court of competent jurisdiction or other governmental body of the United States or any political subdivisions thereof, Recipient shall first (a) notify ISBE of the order and provide a complete copy of such order to ISBE and (b) permit ISBE to seek an appropriate protective order. Recipient shall fully cooperate with ISBE if ISBE wishes to apply to such court for a protective order. Recipient shall only disclose the Confidential Data to the extent necessary and for the purposes of the court or other governmental body. Furthermore, Recipient must comply with the notice requirements of FERPA (34 C.F.R. § 99.31(a)(9)(ii) when and if it is required to disclose any Student Data in accordance with a lawfully issued subpoena or court order. 34 C.F.R. § 99.33(b)(2).F. Recipient must create and maintain a record of any disclosure of Confidential Data made to any other person or entity not already denoted in Exhibits A through E pursuant to this Agreement. The record of disclosure must record the name of any additional person or organization receiving the Confidential Data and their legitimate interest under 34 C.F.R. § 99.31 in requesting or obtaining the Confidential Data. The record must also describe the Confidential Data included within the disclosure by class, school, district, or other appropriate grouping. Upon ISBE's request, Recipient must provide a copy of the record of further disclosures to ISBE. 34 C.F.R. § 99.32(b)(2)(i) and (ii).

F. Nothing in this Agreement may be construed to allow Recipient to maintain, use, disclose, or share the Confidential Data in a manner not allowed by state or federal law or regulation, including but not limited to FERPA (20 U.S.C. § 1232g) and ISSRA (105 ILCS 10/1, et seq.).

G. Recipient will not share Confidential Data with anyone, except those employees and contractors, subcontractors and agents of Recipient with a legitimate interest in the Confidential Data for Audit, Evaluation, or Research, as identified in Exhibit A, Attachments 1 and 2, as may be amended from time to time in accordance with the terms and conditions of this Agreement and consistent with 34 C.F.R. § 99.32(a).

H. Recipient will instruct all persons having access to Confidential Data on the use and confidentiality restrictions set forth in this Agreement and sanctions for unauthorized disclosure and shall require all employees, contractors, subcontractors, or agents of any kind to comply with all applicable provisions of FERPA and other state and federal laws with respect to the Confidential Data. Recipient shall provide executed copies of Exhibit E, Attachment 1 ("Security Pledge for the Use of Confidential Data") for each such person and upon ISBE's request, shall produce a written acknowledgement from all such persons verifying that the instruction required under this Section has occurred.

I. Recipient will not disclose any individually identifiable information or Confidential Data under this Agreement in a manner which could identify an individual student, person, program, school, institution, or district except as authorized by ISBE and applicable law. ISBE expressly authorizes the disclosure of individually identifiable information or Confidential Data by Recipient only if such individually identifiable information or Confidential Data pertains directly to the Research contemplated under this Agreement as described in Exhibit A ("Specifications for Shared Data and Use") subject to the same security provisions in the Agreement. Disclosure includes, without limitation, disclosure of information, research, or analysis in a manner that permits the personal identification of parents and students, as such terms are defined in the FERPA regulations (34 C.F.R Part 99), or individual identification of a person, program, school, institution, or district; and includes, de-identified or aggregate data in cell sizes of less than ten (10) for each category or subcategory of data, and de- identified or aggregate data in cell sizes of more than ten (10) for each category or subcategory that, when disaggregated could lead to indirect disclosure through the disclosure, through the cumulative effects of disclosures, or when combined with other data element(s) in the public domain.

J. Recipient may not re-disclose Student Data to any other person or entity unless permitted or required by law and approved in advance under an amendment to this Agreement. Re-disclosure of Student Data includes, without limitation, disclosure of information, research, or analysis in a manner that permits the personal identification of parents and students, as such terms are defined in the FERPA regulations (34 C.F.R. Part 99); and includes, de-identified or aggregate data in cell sizes of less than ten (10) for each category or subcategory of data, and de-identified or aggregate data in cell sizes of more than ten (10) for each category or subcategory that, when disaggregated could lead to indirect disclosure through the disclosure, through the cumulative effects of disclosures, or when combined with other data element(s) in the public domain.

K. Recipient certifies that it has the capacity to restrict access to the Confidential Data and maintain the security of electronic information, as more fully set forth in Exhibit E ("Data Security Plan"). Recipient shall develop, implement, maintain and use appropriate administrative, technical and physical security measures to preserve the confidentiality, integrity and availability of all electronically maintained or transmitted Confidential Data received from, or on behalf of, ISBE. Recipient acknowledges that the use of unsecured telecommunications, including the Internet or email, to transmit individually identifiable or deducible information derived from the Confidential Data specified in Exhibit A is strictly prohibited.

L. Recipient agrees that all data transferred pursuant to this agreement will be through encrypted transmission mechanisms. These may include but not be limited to secure FTP or web sites using SSL protocols. These measures will be extended by contract to all employees, contractors, subcontractors, or agents that will receive Confidential Data provided by this Agreement and used by Recipient.

M. Recipient will not provide any of the Confidential Data obtained pursuant to this Agreement to any party ineligible to receive data protected by FERPA or prohibited from receiving data from any entity by virtue of a finding under subsections 99.67(c), (d) or (e) of Title 34 of the Code of Federal Regulations. 34 C.F.R. § 99.67 (c), (d) and (e).

N. Recipient agrees to fully report to ISBE within one day of discovery any infraction of the confidentiality provisions and any use or disclosure of Confidential Data not authorized by this Agreement or in writing by ISBE. Recipient's report shall identify: (i) the nature of the unauthorized use or disclosure; (ii) the Confidential Data used or disclosed; (iii) who made the unauthorized use and/or received the unauthorized disclosure; (iv) what Recipient has done or shall do to mitigate any deleterious effect of the unauthorized use or disclosure; and (v) what corrective action Recipient has taken or shall take to prevent future similar unauthorized use and/or disclosure. Recipient shall provide such other information, including a written report, as reasonably requested by ISBE.

O. Recipient agrees that Confidential Data shall not be archived or sent to a records center except as set forth in Attachment 1, Paragraph E.

P. Recipient agrees to secure any and all data received pursuant to this Agreement and agrees to establish, secure and retain records of access and use of all Confidential Data received pursuant to this Agreement. Recipient agrees to allow ISBE on-site inspection and access to all relevant data files and servers to verify data security and usage, as well as audit access, throughout the Term of this Agreement and for a period of three (3) years following the Completion Date, whichever is longer. The three (3) year period shall be extended for the duration of any audit in progress during the Term. No fees shall be assessed for such access, audit, or review, and Recipient agrees to cooperate with ISBE's efforts to verify data security and usage.

Q. Any breach of the security of any Confidential Data provided to any person or entity under this Agreement shall be subject to the terms and provisions of the Personal Information Protection Act (815 ILCS 530/1, et seq.).

R. Recipient represents and agrees that any and all approvals for the research to be conducted using the Confidential Data, where required by law, from the Recipient or the Recipient's Institutional Review Board ("IRB") have been obtained. ISBE may request a copy of any review completed by Recipient or the Recipient's IRB related to the Confidential Data; and Recipient shall provide ISBE with a copy of the requested review within ten (10) working days of ISBE's written request.

S. Recipient may not assign its obligations under this Agreement, or any part of its interest in this Agreement, without the prior written consent of ISBE. Any assignment made without said consent shall be null and void.

T. Recipient recognizes and agrees that the Confidential Data it obtains under this Agreement is the property of ISBE and shall be disposed of or returned to ISBE within ten (10) days, upon ISBE's request. All Confidential Data received pursuant to this Agreement shall be disposed of upon termination, cancellation, expiration, or other conclusion of this Agreement. Disposal means the return of the Confidential Data to ISBE or destruction of the Confidential Data in a means outlined in Exhibit E as directed by ISBE, including purging of all copies from the Recipient's computer systems. Upon disposal of the Confidential Data, Recipient shall provide ISBE with the certificate in Exhibit E, Attachment II. Recipient agrees to require all employees, contractors, subcontractors, or agents of any kind to comply with this provision.

2. Recipient must ensure that any third-party recipient of the Confidential Data working under or in collaboration with Recipient agrees by contractual terms to the provisions of this Agreement for the sharing, disclosure, re-disclosure, use, maintenance, security and destruction of the Confidential Data.

3. The terms and provisions of this Section III shall apply to the use of Confidential Data received by Recipient for so long as Recipient retains the data and shall survive the expiration or earlier termination of this Agreement.

IV. TERM AND TERMINATION

1. **Term.** This Agreement shall become effective on the date of signature of the last signatory to the Agreement and, subject to any earlier termination as provided herein, shall remain in full force and effect through and including December 31, 2017 (the

"Term"). At the sole option of ISBE and subject to the mutual agreement of the Parties, this Agreement may be renewed for one two-year term no cost extension (from January 1, 2018 through December 31, 2018; and from January 1, 2019 through December 31, 2019).

2. **Termination**. This Agreement may be terminated by either Party upon thirty (30) days written notice to the other Party.

3. **Termination for Breach.** Notwithstanding any other provisions to the contrary, this Agreement is subject to immediate cancellation by ISBE for failure of Recipient or its authorized employee, contractor, subcontractor, or agent to adhere to any provision set forth in this Agreement.

4. **Termination upon Unauthorized Data Disclosure.** Notwithstanding any other provisions to the contrary, ISBE may immediately terminate its participation in this Agreement if any Confidential Data disclosed by ISBE to Recipient is used in any manner which violates the terms and provisions of this Agreement, ISSRA and/or FERPA.

5. **Survival.** Notwithstanding any other provision in this Agreement, the terms of this Agreement regarding the use, confidentiality, and secure maintenance of data shall survive the termination of the Agreement and continue in full force and effect.

V. GENERAL PROVISIONS

1. **Amendment**. This Agreement may be amended only by a written instrument signed by the Parties. Notwithstanding the foregoing, any amendment to modify or add to the scope of the research or data to be shared must be signed by Recipient and the Illinois State Superintendent of Education and ISBE General Counsel and shall be in the form as set forth in Exhibit B (the "Form for Amendment of Research Scope or Shared Data").

2. **Reservation of Data Release**. ISBE reserves the right to refuse any data request involving individually identifiable information data or school/program/institution/district level data. However, nothing herein shall prohibit individuals or entities from releasing data pertaining to themselves or their own school, program, institution or district.

3. **Comment and Approval Period**. ISBE shall be provided for its review, any and all research and other reports produced using its data. The Recipient will provide ISBE with one electronic and at least one paper copy of a final draft and all final versions of all approved reports to be released, along with other documents associated with any of the Specifications, as set forth in Exhibit A. ISBE expressly reserves the right to review, comment, and approve any use of the data shared or collected pursuant to this Agreement before its public release. The period of such review will be 60 days from ISBE's receipt of the material to be publicly released. After the 60 day review period has lapsed, if ISBE has not approved the use in writing, the Recipient may not release the material publicly with ISBE's data included, but shall remove all ISBE data from the proposed publication and associated research analysis and provide a copy of the proposed publication, sans ISBE data, to ISBE prior to final publication. ISBE reserves the right to demand, and,

if requested, Recipient shall include in any material to be publicly released that includes the use of ISBE data, a rejoinder to be provided in writing by ISBE.

4. **Public Announcements**. All media releases and public announcements by either Party relating to this Agreement, the Research Project, or the Specifications shall be coordinated with and approved in writing by ISBE.

5. **Authorized Representatives**. The following persons are authorized to approve an amendment to this Agreement on behalf of the Parties (each is an "Authorized Representative;" collectively, the "Approval Representatives"):

   For ISBE: The State Superintendent of Education, or designee, with form approved by the General Counsel, or designee(s);

   For Recipient: Kathleen Parks, Vice President.

6. **Notices.** All notices or other correspondence required to be given pursuant to this Agreement shall be sent by mail or delivered to the Parties' Approval Representatives at the following addresses:

   For ISBE:

   General Counsel
   Illinois State Board of Education
   100 North First Street
   Springfield, Illinois 62777
   Fax: (217) 524-3911

   For Recipient:

   Kathleen Parks
   NORC at the University of Chicago
   1155 East 60th Street, 3rd FL
   Chicago, IL 60637

   In case of an emergency or when immediate assistance is needed:

   The person to contact on behalf of ISBE is:

   Marjurie Ribeiro
   Telephone (312) 814-7239

   The person to contact on behalf of Recipient is:

   Eric Hedberg, Senior Research Scientist
   Telephone (773) 909-6801

7. **Entirety.** This Agreement, together with the Exhibits attached hereto, constitutes the entire Agreement among the Parties with respect to the subject matter hereof and supersedes any other negotiations, agreements, or communications, whether written or oral, that have been made by any Party.

8. **Severability.** If any provision of this Agreement shall be held invalid, illegal, or unenforceable, such provision shall be deemed deleted from this Agreement and replaced by a valid and enforceable provision which so far as possible achieves the Parties' intent in agreeing to the original provision. The remaining provisions of this Agreement shall continue in full force and effect.

9. **Governing Law.** This Agreement shall be governed by and construed in accordance with the laws of the State of Illinois. Any claim against the State or a State agency arising out of this Agreement must be filed exclusively with the Illinois Court of Claims (705 ILCS 505/1) when said claim is within the jurisdiction of the Court of Claims.

10. **Records**. Books and records, including information stored in databases or other computer systems, shall be maintained by Recipient for a period of three (3) years from the later of the Term of this Agreement or the Completion Date and by any subcontractor for a period of three (3) years from the later of the date of the final payment under the subcontract or completion of the subcontract. The three (3) year period shall be extended for the duration of any audit in progress during the Term. Books and records required to be maintained under this section shall be available for review or audit by representatives of ISBE, the Auditor General, and other governmental entities with monitoring authority upon reasonable notice and during normal business hours. Recipient and its employees, contractors, subcontractors and agents shall cooperate fully with any such audit. Recipient and its employees, contractors, subcontractors and agents shall not impose a charge for audit or examination of Recipient's or Recipient's contractor's, or subcontractor's books and records.

11. **Hold Harmless**. To the fullest extent allowed by Illinois law, the Recipient agrees to defend, indemnify and hold harmless ISBE, its officers, staff, employees, and agents against any and all claims, suits, damages and causes of action arising out of or in any way related to the activities to be carried out pursuant to the obligations of this Agreement, including but not limited to, the use or disclosure by Recipient, its employees, contractors or agents, of any information received from or through ISBE pursuant to the terms of this Agreement.

12. **Injunctive Relief.** Recipient agrees that an impending or existing violation of any provision of this Agreement would cause ISBE irreparable injury for which it would have no adequate remedy at law and that ISBE shall be entitled to seek immediate injunctive relief prohibiting such violation, in addition to any other rights and remedies available to it.

13. **Authority to Execute.** Each Party represents and warrants to the other Party that this Agreement has been duly authorized, executed and delivered by and on behalf of each such Party and constitutes the legal, valid and binding agreement of said Party.

14. **Counterparts.** This Agreement may be executed in several counterparts, each of which shall be an original and all of which shall constitute one and the same instrument. Facsimiles of signatures shall constitute acceptable, binding signatures for purposes of this Agreement.

15. **Recitals and Exhibits.** The recitals in Section I and the following Exhibits are hereby incorporated by reference and expressly made a part of this Agreement.

EXHIBIT A – SPECIFICATIONS FOR SHARED DATA AND USE (INCLUDING ATTACHMENTS 1, 2, 3 AND 4)

EXHIBIT B – FORM FOR AMENDMENT OF RESEARCH SCOPE OR SHARED DATA

EXHIBIT C – FORM FOR AMENDMENT OF EXHIBIT A, ATTACHMENT 1

EXHIBIT D– FORM FOR AMENDMENT OF EXHIBIT A, ATTACHMENT 2

EXHIBIT E – DATA SECURITY PLAN (INCLUDING ATTACHMENTS 1 AND 2)

IN WITNESS WHEREOF, the Parties have executed this Agreement on the dates set forth below.

[redacted] 6/2/15
Signature Date
Division Administrator,

Illinois State Board of Education

DAVID SMALLEY
Print Name

[redacted] 6.8.15
Signature Date
Acting General Counsel
Illinois State Board of Education

Marcilene Dutton
Print Name

[redacted] 6/10/15
Signature Date
State Superintendent of Education
Illinois State Board of Education

Anthony M. Smith
Print Name

[redacted] 5-14-2015
Signature Date
Vice President

National Opinion Research Center, Recipient

Kathleen E. Parks
Print Name

DATA SHARE AND USE AGREEMENT

EXHIBIT A: SPECIFICATIONS FOR SHARED DATA AND USE

RESEARCH STUDY: Advancing State Specific Design Parameters for Designing Better Evaluation Studies

A. DESCRIPTION OF RESEARCH (*including specific research questions*):

This project seeks to use entire state data systems to estimate intraclass correlations, R2 values representing covariate effectiveness, and variance components representing heterogeneity of effects for use in designing multi-level studies in education. This project advances previous work in this field for two reasons. Use of state data systems allows for unprecedented ability to estimate variance components and intraclass correlations at several levels of analysis inclusive of district, school, and classrooms. Second, access to comprehensive data from states allows for estimates of heterogeneity of important effects across district, school, and classes, providing the research community with a better sense of the heterogeneity of program effects that might be reasonable for specific grades and subjects.

The following research questions will be asked:

1. What is the set of grade and subject specific variance components, and intraclass correlations, for cross- sectional achievement at the district, school, classroom, and student, levels?
2. What is the set of grade and subject specific variance components, and intraclass correlations, for achievement gains at the district, school, classroom, and student, levels?
3. How do variance components, and intraclass correlations, differ for each design models?
4. How do the estimated variance components, and intraclass correlations, differ across different school context?
5. What is the set of grade and subject specific effect sizes for a single year increase in achievement?
6. How do these effect sizes differ when we control for student characteristics?
7. How do our effect sizes differ across different school contexts?
8. How do the estimated effect sizes vary across district, school, and classroom contexts?

B. SPECIFIC DATA REQUESTED *(including format and data elements):*

The following data are requested in Excel format: two consecutive years (2012-2013 and 2013-2014) of math and language arts/reading assessments, gender, race/ethnicity, English learner status, a measure of poverty (e.g., eligibility for free or reduced price lunch), disability status, school identifiers and district identifiers.

C. HOW DATA WILL BE UTILIZED *(Including Attachment I and II):*

Data will be used for research purposes. Please see attached project proposal (pg. 30).

D. REPORT:

This project will produce academic papers. NORC will share with ISBE all reports and academic papers utilizing or based on ISBE Confidential Data, as requested.

E. TIMELINE FOR RESEARCH AND RETURN OR DESTRUCTION OF CONFIDENTIAL DATA: *(Provide a basic timeline for the research and anticipated completion date, with dates for return or destruction of confidential data.)*

Recipient's proposal contemplates a 36 month project with a completion date of December 31, 2017.

Destruction of Confidential Data will occur in compliance with this Agreement, six (6) months following the completion of the Research and Evaluation and consistent with IRB specifications (20 CFR § 46.115)

DATA SHARE AND USE AGREEMENT

EXHIBIT A: SPECIFICATIONS FOR SHARED DATA AND USE

ATTACHMENT 1

Name, position and legitimate interest in the Confidential Data, of all officials and employees authorized to request, receive, and obtain information, including Confidential Data, from ISBE or Recipient under this Agreement.

Name	**Position and Legitimate Interest**
Eric Hedberg	Senior Research Scientist, Research

DATA SHARE AND USE AGREEMENT

EXHIBIT A: SPECIFICATIONS FOR SHARED DATA AND USE

ATTACHMENT 2

Name of each contractor, subcontractor, or agent and his/her position, legitimate interest in the Confidential Data, organizational affiliation, address, telephone number and facsimile number who will request, receive, or obtain information, including Confidential Data, from Recipient under this Agreement.

N/A

Name:	
Title:	
Organization:	
Street Address:	
City/State/ZIP Code:	
Voice:	
Fax:	
E-Mail:	
Legitimate Interest in Confidential Data	
Name:	
Title:	
Organization:	
Street Address:	
City/State/ZIP Code:	
Voice:	
Fax:	
E-Mail:	
Legitimate Interest in Confidential Data	

DATA SHARE AND USE AGREEMENT

EXHIBIT A: SPECIFICATIONS FOR SHARED DATA AND USE

ATTACHMENT 3

DATA ELEMENTS TABLES

1. Assess Test Code table (tbAssessTest)
8. ELL Entry Status code Table (tbSISELLEntryStatus)
10. ELL Exit Type code Table (tbSISELLExitType)
11. English Language Learner ELL Table (tbSISELL)
15. Grade Level Code Table (tbSISGrdLvl)
16. Homeless Table (tbSISHomeless)
18. Native Language Code Table (tbSISNativeLang)
20. Race Code Table (tbSISRace)
22. SIS Assessment Demographics table (tbSISAssesCorrections)
23. SIS Assessment Scores Table (tbSISAssessTestingScores)
24. SIS Demographic (tbSISStu)(keyed)

DATA SHARE AND USE AGREEMENT

EXHIBIT A: SPECIFICATIONS FOR SHARED DATA AND USE

ATTACHMENT 4

Assesment Test Code Table
chrAssessTestCode
vchrAssessTestDesc

ELL Entry Status code Table intELLEntryStatusId
chrELLEntryStatusCode
vchrELLEntryStatusDesc

ELL Exit Status code Table
intELLExitTypeId
chrELLExitTypeCode
vchrELLExitTypeDesc

ELL Table
intELLId
intNativeLangId
chrDateFirstEnrolled

intELLENtryStatusId
chrELLExitDate
bitTitleIIIStatus
bitImmigrantEdProg
bitMigrant
bitSpEdIEP

Grade Level Code Table
intGrdLvlId
chrGrdLvlCode
vchrGrdLvlDesc

Homeless Table
intEnrlID
bitLEPInd
bitIEPInd
bitMigrant

Native Language Code Table
intNatLangId
chrNatLangCode
vchrNatLangDesc

Race Code Table
intRaceId
chrRaceCode
vchrRaceDesc

Assessment Demographics Table intsystemsuid
intFY
chrRace
chrGender
chrHomelessInd
chrMigrant
chrLEPInd
chrIEPInd
chrFRLInd

SIS Assessment Scores Table
intScoresId
intAssesCorrId
intFY
intAssesSeqNum
chrGradeOfTest
chrAssessTestCode
chrGradeOfTestRdng
chrGradeOfTestMath
chrRdngScaleScore
chrRdngPerfLvl
chrMathScaleScore
chrMathPerfLvl

SIS Student Record Table intSystemStuId
chrGender
intNativeLangId
bitHomelessInd
intRaceId
bitMigrant
bitLEPInd
bitIEPInd
bitFRLInd
bitSpecEdInd
bitSESInd
bitTitleOneInd
chrTitleOneCode

DATA SHARE AND USE AGREEMENT

EXHIBIT B

FORM FOR AMENDMENT OF RESEARCH SCOPE OR SHARED DATA

AMENDMENT # ___ TO EXHIBIT A OF THE DATA SHARE AND USE AGREEMENT BETWEEN THE ILLINOIS STATE BOARD OF EDUCATION AND the National Opinion Research Center ("NORC")

Dated__________________

A. **DESCRIPTION OF RESEARCH** (*including additional or modified specific research questions*):

B. **SPECIFIC DATA REQUESTED** (*including format with any additionally requested data elements*):

C. **HOW DATA WILL BE UTILIZED** (*Including Modified Attachment(s) I and II*):

D. **REPORT:**

E. **TIMELINE FOR RESEARCH AND RETURN OR DESTRUCTION OF CONFIDENTIAL DATA:** *(Provide a modified basic timeline for the research and anticipated completion date, with dates for return or destruction of confidential data.)*

IN WITNESS WHEREOF, the Parties have executed this Amendment on the dates set forth below.

State Superintendent of Education or Designee

National Opinion Research Center

Title

Vice President, Academic Research Center
Title

Print Name

Kathleen E. Parks
Print Name

Signature Date

[redacted] 5-14-2015
Date

Approved:

General Counsel or Designee
Illinois State Board of Education

Title

Print Name

Signature Date

DATA SHARE AND USE AGREEMENT

EXHIBIT C

FORM FOR AMENDMENT OF EXHIBIT A, ATTACHMENT 1

AMENDMENT # , AMENDING ATTACHMENT 1 OF EXHIBIT A OF THE DATA SHARE AND USE AGREEMENT BETWEEN THE ILLINOIS STATE BOARD OF EDUCATION AND the National Opinion Research Center ("NORC")

Dated ____________

Name position and legitimate interest in the Confidential Data, of all officials and employees authorized to request, receive, and obtain information, including Confidential Data, from ISBE or Recipient under this Agreement.

Name	**Position and Legitimate Interest**

Submitted by:

[Insert Recipient]

DATA SHARE AND USE AGREEMENT

EXHIBIT D

FORM FOR AMENDMENT OF EXHIBIT A, ATTACHMENT 2

AMENDMENT # , AMENDING ATTACHMENT 2 OF EXHIBIT A OF THE DATA SHARE AND USE AGREEMENT BETWEEN THE BETWEEN THE ILLINOIS STATE BOARD OF EDUCATION AND the National Opinion Research Center ("NORC")

Dated ____________

Name of each contractor, subcontractor, or agent and his/her position, legitimate interest in the Confidential Information, organizational affiliation, address, telephone number and facsimile number who will request, receive, or obtain information, including Confidential Data, from Recipient under this Agreement.

Name:

Title:

Organization:

Street Address:

City/State/ZIP Code:

Voice:

Fax:

E-Mail:

Legitimate Interest in Confidential Data

Submitted by: Date:

[Insert Recipient]

DATA SHARE AND USE AGREEMENT

EXHIBIT E

DATA SECURITY PLAN

Capitalized terms used herein, unless otherwise defined in this Data Security Plan ("Plan"), shall have the respective meanings assigned to them in the Agreement.

For and in consideration of the agreement to provide the Recipient with the Confidential Data and the mutual covenants contained herein and in the Agreement, Recipient and ISBE hereby agree as follows:

A. GENERAL INFORMATION

a. The Confidential Data will only be stored in an appropriate manner as defined below.

b. Only one complete copy of the Confidential Data is permitted to be maintained by Recipient; however, time-delimited temporary data analysis files may be created. Any temporary data file(s) and subsets of the original data set will be considered Confidential Data and subject to the terms and conditions of this agreement.

B. PROTECTION OF DATA

All Confidential Data shall be stored in a secure environment physically located in the continental United States with access limited to the least number of staff needed to complete the purpose of this Agreement.

Recipient agrees to store data on one or more of the following media and protect the data as described:

a. Data stored on local workstation hard disks. Access to the data will be restricted to authorized users by requiring logon to the local workstation using a unique user ID and complex password or other authentication mechanisms which provide equal or greater security, such as biometrics or smart cards. If the workstation is located in an unsecured physical location the hard drive must have encryption to protect the Confidential Data in the event the device is stolen.

b. Data stored on hard disks mounted on network servers and made available through shared folders. Access to the data will be restricted to authorized users through the use of access control lists which will grant access only after the authorized user has authenticated to the network using a unique user ID and complex password or other authentication mechanisms which provide equal or greater security, such as biometrics or smart cards.

Data on disks mounted to such servers must be located in an area which is accessible only to authorized personnel, with access controlled through use of a key, card key, combination lock, or comparable mechanism. Backup copies for DR purposes must be encrypted if recorded to removable media.

c. Paper documents. Any paper records must be protected by storing the records in a secure area which is only accessible to authorized individuals. When not in use, such records must be stored in a locked container, such as a file cabinet, locking drawer, or safe, to which only authorized persons have access.

d. Access via remote terminal/workstation over the Public Internet. Remote data access is prohibited unless Recipient requests remote access and ISBE authorizes remote access as part of this agreement. If requesting remote access the Recipient will include the safeguards in place to secure the receipt and transmission of data.

e. Confidential Data shall not be stored by Recipient on portable devices or media which include but are not limited to laptops, tablets, handhelds/PDAs, Ultramobile PCs, optical discs, CDs, DVDs, Blu-Rays, removable storage and flash memory devices unless specifically requested by the Recipient and authorized within this Agreement. The request must include methods for encrypting the data, controlling access to the data and physically protecting the device(s) containing the data.

C. DATA SEGREGATION

a. Confidential Data must be segregated or otherwise distinguishable from non-Confidential Data. This is to ensure that when no longer needed by the Recipient, all Confidential Data can be identified for return or destruction. It also aids in determining whether Confidential Data has or may have been compromised in the event of a security breach.

b. Confidential Data shall be stored in one of the following methods:

   i. Confidential Data will be kept on media (e.g. hard disk, optical disc, tape, etc.) which will contain no non- Confidential Data; or
   ii. Confidential Data will be stored in a logical container on electronic media, such as a partition or folder dedicated to confidential data; or,
   iii. Confidential Data will be stored in a database which will contain no non-Confidential Data; or,
   iv. Confidential Data will be stored within a database and will be distinguishable from non- Confidential Data by the value of a specific field or fields within database records; or
   v. When it is not feasible or practical to segregate Confidential Data from non-Confidential Data, then both the confidential data and the non-confidential data with which it is commingled must be protected as described in this Agreement.

c. If the Recipient or its agents detect a compromise or potential compromise in the IT security for this data such that personal information may have been accessed or disclosed without proper authorization, Recipient shall give notice to ISBE as outlined previously in this agreement.

D. DISPOSITION OF DATA

a. Upon termination of the agreement, Recipient shall dispose of the data received along with backup copies and any temporary or permanent work files that contain confidential data and provide written notification of disposal. Failure to do so may prevent data sharing agreements with the organization in the future.

b. Upon the destruction of the confidential data, the Recipient shall complete Attachment 2 of this Exhibit, Certification of Data Disposition, and submit it to the ISBE authorized representative within fifteen (15) days of the date of disposal.

c. Acceptable destruction methods for various types of media include:

1. For paper documents containing confidential or sensitive information, a contract with a recycling firm to recycle confidential documents is acceptable, provided the contract ensures that the confidentiality of the data will be protected. Such documents may also be destroyed by on-site shredding, pulping, or incineration.
2. For paper documents containing Confidential Data requiring special handling, recycling is not an option. These documents must be destroyed by on-site shredding, pulping, or incineration.
3. If confidential or sensitive information has been contained on optical discs (e.g. CDs, DVDs, Blu-ray), the data recipient shall either destroy by incineration the disc(s), shredding the discs, or completely deface the readable surface with a coarse abrasive.
4. If confidential or sensitive information has been stored on magnetic tape(s), the data recipient shall destroy the data by degaussing, incinerating or crosscut shredding.
5. If data has been stored on server or workstation data hard drives or similar media, the data recipient shall destroy the data by using a "wipe" utility which will overwrite the data at least three (3) times using either random or single character data, degaussing sufficiently to ensure that the data cannot be reconstructed, or physically destroying disk(s).

6. If data has been stored on removable media (e.g. floppies, USB flash drives, portable hard disks, or similar disks), the data recipient shall destroy the data by using a "wipe" utility which will overwrite the data at least three (3) times using either random or single character data, degaussing sufficiently to ensure that the data cannot be reconstructed, or physically destroying disk(s).

DATA SECURITY PLAN

ATTACHMENT 1

Security Pledge for the Use of Confidential Data

I, Eric Hedberg through my involvement with and work with The National Opinion Research Center ("NORC"), will have access to confidential data collected by the Illinois State Board of Education ("ISBE"). By virtue of my affiliation with the research projects being lead by Eric Hedberg, Senior Research Scientist, I have access to confidential information and use of (a) data about students that is considered personal and private under the Family Educational Rights and Privacy Act (20 U.S.C.§ 1232g) and the Illinois School Students Records Act (105 ILCS 10/1, *et seq*.), (b) data about teachers that is generally perceived as personal and private, and (c) confidential information about programs, schools, institutions, and districts. I understand that access to this confidential information and data carries with it the responsibility to: (a) guard against unauthorized use, (b) abide by all security parameters, requirements and guidelines instituted by the NORC in conjunction with its obligations regarding the confidential data, and (c) abide by the Data Security Plan entered into between ISBE and the NORC. To treat information as confidential means not to divulge it to anyone who is not a project member or to cause it to be accessible to anyone who is not a project member. I understand that the use of unsecured telecommunications, including the Internet, to transmit individually identifiable or deducible information derived from the data is strictly prohibited, and agree that all data transmissions must be approved by Kathleen Parks prior to transmission and must be encrypted and provided through a secure FTP site. Anything not specifically named as "public information" is considered confidential.

Each person using data collected by ISBE is reminded that disclosing confidential information directly or allowing non-authorized access to such information may subject that individual to criminal prosecution and/or civil recovery.

I agree to fulfill my responsibilities on this project in accordance with the following guidelines:

1. I agree not to permit non-project personnel access to these sensitive data, either electronically or in hard copy.

2. I agree not to attempt to disclose individuals, families, households, programs, schools, districts, or institutions.

3. I agree that in the event an identity of an individual, family, household, program, school, district, or institution is discovered inadvertently, I will (a) make no use of this knowledge, (b) advise the NORC of the incident, who will report *it* to ISBE, (c)

safeguard or destroy the information as directed by the NORC after consultation with ISBE, and (d) not inform any other person of the discovered identity.

[redacted]
Signature

5-6-2015
Date

DATA SHARE AND USE AGREEMENT

EXHIBIT A: SPECIFICATIONS FOR SHARED DATA AND USE

C. HOW DATA WILL BE UTILIZED (continued from pg. 14)

Advancing State-specific Design Parameters for Designing Better Evaluation Studies

Project Narrative. This project seeks to use entire state data systems to estimate intraclass correlations, R^2 values representing covariate effectiveness, and variance components representing heterogeneity of effects for use in designing multi-level studies in education. This project advances previous work in this field for two reasons. Use of state data systems allows for unprecedented ability to estimate variance components and intraclass correlations at several levels of analysis inclusive of district, school, and classrooms. Second, access to comprehensive data from states allows for estimates of heterogeneity of important effects across district, school, and classes, providing the research community with a better sense of the heterogeneity of program effects that might be reasonable for specific grades and subjects. This project will produce reports to IES, academic papers, and for participating states, workshops and technical assistance in their evaluation designs.

Significance. Randomized experiments are the most reliable instruments for assessing the causal effects of educational interventions, products, and services. The vast majority of educational filed experiments fall into one of two major types: cluster randomized designs that assign entire intact groups to treatments or (generalized) randomized block designs that assign individuals or groups to treatments within larger intact groups or blocks (Spybrook and Raudenbush, 2009). Randomized block designs are sometimes called multi-site or matched designs.

Cluster Randomized Designs. In educational settings randomization of individuals to treatments may be theoretically impossible (e. g., when the treatment is applied to a whole school such as school wide positive behavior support) or infeasible for practical or political reasons. Even when it is possible, individual assignment may be methodologically unwise because the close proximity of individuals receiving (or providing) different treatment can lead to treatment diffusion and contamination effects that can compromise the internal validity of the experiment.

For these reasons, many educational experiments use designs that involve the random assignment of entire pre-existing groups to treatments. Studies might assign whole classrooms, schools, or school districts to treatments. The groups typically used in education experiments are not themselves composed at random, but exhibit intraclass correlation structure. That is, the individuals in the same group tend to be more alike than individuals in different groups, which implies a kind of dependence among individuals in groups that statisticians since R. A. Fisher have called intraclass correlation. In sample survey work, the same phenomenon would be described by saying that the intact groups were statistical clusters and educational experiments assigned clusters to treatments, which is why group randomized designs are sometimes called cluster randomized designs.

Populations of students in education often exhibit multiple levels of clustering, because students are clustered within classes, classes are clustered within schools, and schools are clustered within school districts. If (but only if) the sampling of individuals exploits this population structure, there may be a multilevel intraclass correlation structure. For example individuals in the same school, but different classes may tend to be more alike than individuals in different schools, and one intraclass correlation parameter might describe this phenomenon. However, two individuals in the same classroom might tend to be more alike than two individuals in the same school (but in different classrooms), and a second intraclass correlation parameter might describe this phenomenon. Thus, in this three level situation (individuals, classrooms, and schools as the levels of aggregation), two intraclass correlation parameters are necessary to describe the clustering structure. In a four level structure (individuals, classrooms, schools, and school districts as the levels of aggregation), three intraclass correlation parameters would be needed to describe the clustering structure.

Randomized Block Designs. It is often desirable to increase precision and statistical power in designs by matching units assigned to different treatments within naturally occurring units at a higher level of aggregation (called blocks in experimental design), like schools within districts, classroom or teachers within schools, or individuals within classrooms. Of course, randomized block designs involve assigning units within the same larger aggregate unit to different treatments, and consequently there is some potential for contamination between treatments. However there is often little reason to expect substantial contamination and recent theoretical work has demonstrated that even if there is some contamination, randomized block designs may still have substantially more power than cluster randomized designs involving the same sample sizes (Rhoads, 2011). While cluster randomized designs are more common in educational research, randomized block designs are also rather common (see Spybrook and Raudenbush, 2009) and may become more so as resources for funding research studies in education become more constrained.

As in the case of cluster randomized designs, clustered sampling designs used in randomized block designs lead to intraclass correlation structures that have implications for design sensitivity. However, in the case of randomized block designs, it is not merely the variation of cluster (or subcluster) means that has implications for design sensitivity, but also variation of treatment effects across clusters (or subclusters). For example, in a three level design that assigns treatments to classes within schools, a treatment effect is (in principle) estimable within each school, and the variation of the treatment effects across schools is an important design parameter. Moreover, treatment effect variation has potentially the largest effect on sensitivity in randomized block designs, analogous to that of the cluster level intraclass correlation in the cluster randomized design.

Statistical Power. It is universally acknowledged that wise design of experiments should include attention to statistical power and the closely related concept of precision of estimates. Statistical power is the probability that the experiment will detect a treatment effect (by rejecting the corresponding null hypothesis in a statistical significance test). A primary design consideration should be to ensure that the experiment has sufficient statistical power to detect the treatment effect that is anticipated, or the smallest treatment effect that is deemed to be substantively meaningful. In individually randomized designs, statistical power depends on the level of statistical significance chosen, the effect size (in standardized form), and the sample size, but power computations are somewhat more complex in cluster randomized and randomized block designs.

Statistical Power in Cluster Randomized Designs. In cluster randomized experiments without covariates, the statistical power depends on how the sample size (the total number of individuals) is allocated among the clustering structure (that is how many units there are at each level) and on the intraclass correlation structure (see, e.g., Hedges & Rhoads, 2010). Two experiments with the same total sample size can have dramatically different power to detect a given effect size, even assuming a constant intraclass correlation.

Covariates are often used to increase power and precision in education experiments. If there are covariates, the power in an individually randomized experiment also depends on the explanatory power of the covariates. In group randomized experiments, power also depends on the explanatory power of the covariates at each level of the design. For example in a two-level experiment (e.g., one assigning schools to treatments with individuals within schools) analyzed using covariates at both group and individual levels, the statistical power will depend on both the explanatory power of the covariates at the group (school) level and, to a lesser extent, on the explanatory power of the covariates at the individual level.

Specifically, the statistical power is determined by the noncentrality parameter associated with each design. As an example, consider the balanced three level design that assigns m schools to each treatment but in which p classrooms of size n students are nested within each school (for a total sample size of $2mpn$). The statistical power depends on the noncentrality parameter

$$\lambda = \delta\sqrt{\frac{mpn/2}{pn\left(1-R_S^2\right)\rho_S + n\left(1-R_C^2\right)\rho_C + \left(1-R_I^2\right)\left(1-\rho_S-\rho_C\right)}},$$

where δ is the effect size (standardized mean difference standardizing by the total standard deviation within treatment groups, see Hedges, 2007 or Hedges, 2009), ρ_S is the school level intraclass correlation, ρ_C is the classroom level intraclass correlation, R_S^2 is the proportion of variation at the school level explained by the covariates, R_C^2 is the proportion of variation at the classroom level explained by the covariates, and R_I^2 is the proportion of variation at the individual level explained by the covariates. Different software programs can require different particular configurations of input parameters, but any valid power analysis requires all of the parameters necessary to compute λ.

Statistical Power in Randomized Block Designs. In randomized block experiments without covariates, the statistical power depends on how the sample size (the total number of individuals) is allocated among the clustering structure (that is how many units there are at each level) on the intraclass correlation structure, and on the heterogeneity of treatment effects (see, e.g., Hedges & Rhoads, 2010). Two experiments with the same total sample size can have dramatically different power to detect a given effect size, depending upon how the sample is allocated.

Covariates may also be used to increase statistical power in randomized block designs, but in such designs the power may depend on the effectiveness of the covariates in explaining variation in treatment effects across blocks. For example in a two-level randomized block experiment (e.g., one assigning individuals to treatments within schools) analyzed using covariates at both group and individual levels, the statistical power will depend on both the explanatory power of the covariates in explaining variation in treatment effects at the group (school) level and, to a lesser extent, on the effectiveness of covariates in explaining scores at the individual level.

As in the case of cluster randomized designs, the statistical power is determined by the noncentrality parameter associated with each design. As an example, consider the balanced three level design that assigns p classrooms with n students to each treatment within a total of m schools (for a total sample size of $2mpn$), but in which classrooms are nested within schools. The statistical power depends on the noncentrality parameter

$$\lambda = \delta\sqrt{\frac{mpn}{2\left(1-R_I^2\right)\left(1-\rho_S-\rho_C\right)+pn\left(1-Q_S^2\right)\omega_S\rho_S+2R_C^2 n\rho_C}}$$

where δ, ρ_S, ρ_C, R_C^2, and R_I^2 are defined as in the cluster randomized experiment, the Q_S^2 is the proportion of variance explained by the level 3 covariates in the treatment effect variance across schools, and ω_S is a heterogeneity parameter (the ratio of treatment effect variance component across schools to the variance component of school means). Different software programs can require different particular configurations of input parameters, but any valid power analysis requires all of the parameters necessary to compute λ.

In randomized block designs with different numbers of levels, or where the treatment is assigned at a different level, there is a treatment heterogeneity parameter at every level above the one at which the treatment is assigned (see, e.g., Hedges and Rhoads, 2010). For example, if treatments are assigned within classrooms in a three level design, there is a treatment heterogeneity parameter reflecting variation of treatment effects across classes within schools, and one reflecting the variation of school average treatment effects across schools.

Fixed Blocking Variables. Note That If Higher Level Aggregate Units Are Treated As Fixed Blocks Or the treatment effect is assumed to be identical in every block, then the design can be seen as a cluster randomized design with the blocks functioning as categorical covariates. The former assumption is not always desirable (e.g., if there is a desire to generalize to other blocks not in the experiment). The latter assumption is not always tenable (e.g., because there is no reason to assume treatment effects will be the same across blocks). If these assumptions are not met, analyses of design sensitivity (power and precision) will only be correct if the design is treated as a randomized block design.

Researchers often use fixed blocking variables at the highest aggregate level of the design to increase power and precision (see, e. g., Raudenbush, Martinez, and Spybrook, 2007). For example, in a design that (conceptually) has three levels districts, schools, and individuals, one might block by districts and assign schools to treatments. To compute statistical power or precision in such a design, the proportion of total variance at the school level (level 2) that is accounted for by the blocking variable (districts) is required, and consequently software like Optimal Design requires this parameter. Because blocking by district accounts of all of the variation at the district level, the proportion of school level (level 2) variance accounted for by the blocking factor (districts) is $\rho_D/(\rho_D + \rho_S)$

where ρ_D and ρ_S are the district and school level intraclass correlations, respectively. Similar identities can be used to obtain the proportion of variance accounted for by blocking variables that might be used at other levels of the design. Therefore the intraclass correlations at various levels can be used to determine the variance accounted for by blocking variables.

Precision of Treatment Effect Estimates. Some analysts prefer to focus on the precision of the estimate of the treatment effect from a design rather than statistical power (this is in keeping with a preference for interval estimation over hypothesis testing). The form of the precision of estimates of the treatment effect is quite similar to the noncentrality parameter that determines statistical power. Table 2 gives the form of the precision of treatment effect estimates from two, three, and four level cluster randomized designs. As an example, consider the same three level cluster randomized design considered above in connection with power analysis. The precision of the estimated treatment effect is

$$\sigma\sqrt{\frac{pn\left(1-R_S^2\right)\rho_S + n\left(1-R_C^2\right)\rho_C + \left(1-R_I^2\right)\left(1-\rho_S-\rho_C\right)}{mpn/2}},$$

where σ is the total standard deviation within treatment groups and the other symbols are defined in the same way as above (see, e. g., Hedges and Rhoads, 2010). Similarly the precision of the estimated treatment effect in the randomized block design considered above is

$$\sigma\sqrt{\frac{2\left(1-R_I^2\right)\left(1-\rho_S-\rho_C\right) + pn\left(1-Q_S^2\right)\omega_S\rho_S + 2R_C^2 n\rho_C}{mpn}},$$

when the symbols are defined in the same way as above (see, e. g., Hedges and Rhoads, 2010).

Optimal Design. In addition to considerations of statistical power and precision of estimates of treatment effects, there is also the question of efficiency. In group randomized (multilevel) designs, many different allocations of sample numbers to levels of the design can yield the same statistical power and precision, and some of these involve more efficient use of resources than others. For example if there is a cost associated with each school in a design (for recruitment travel of research staff to the school, etc.) and a smaller cost associated with each student (for testing, etc.) then a design with more students and fewer schools that yields the same power would be more efficient. This can be formalized in terms of costs per unit at each level of the design, and there is an optimal allocation of units at each level of the design: The allocation that achieves the desired power or precision at the smallest cost. As an example, consider the same three level cluster randomized design considered above in connection with power analysis. The optimal number n_O of individuals for each level 2 unit (e.g., classroom) and the optimal number p_O of level 2 units per level 3 unit are given by

$$n_O = \sqrt{\left(\frac{c_C}{c_I}\right)\left(\frac{\left(1-R_I^2\right)\left(1-\rho_C-\rho_S\right)}{\left(1-R_C^2\right)\rho_C}\right)}, \quad p_O = \sqrt{\left(\frac{c_S}{c_C}\right)\left(\frac{\left(1-R_C^2\right)\rho_2}{\left(1-R_S^2\right)\rho_3}\right)}$$

where c_S is the cost per level 3 unit (e.g., school), c_C is the cost for each level 2 unit (e.g., classroom), and c_I is the cost per each level 1 unit (e.g., individual) and all other symbols are as above (see Konstantopoulos, 2009). Of course many considerations enter into planning research designs, but cost efficiency is always a concern and optimal allocation is often a very useful in planning research designs. Analogous formulas are available for randomized block designs.

Minimum Detectable Effect Size. The minimum detectable effect size is the smallest effect size for which the design has some prespecified power (e.g., 80%) (Bloom, 1995). The minimum detectable effect size is a convenient way to express the sensitivity of designs and is often used in planning designs. The minimum detectable effect size depends on the same parameters as does the precision.

Sound design of experiments involving group randomization requires accurate information on the design parameters that influence statistical power, precision, and optimal design. In essence it requires accurate information about the intraclass correlation structure and the effectiveness of covariates in explaining variation at each level of the research design where they will be used. It also requires information about the variation of treatment effects in randomized block designs.

The statistical power, precision of estimates of treatment effects, the most efficient allocation of sample between levels, and the minimum detectable effect size all depend on the intraclass correlation structure and (if covariates are used) the effectiveness of the covariates in explaining variation at each level where they are used. In a randomized block design, they also depend on the variation of treatment effects. Accurate information about these design parameters is essential for well-designed experiments.

This proposal is for the development of estimates from several state databases, for making these available on an accessible website, and for the development of software tools and procedures to enhance the state's capacity to update and compile estimates to enhance their own research capacity.

Sources of Information About Design Parameters. There are essentially four sources of information about design parameters: experiments that have already been conducted, values obtained from a few large school districts, surveys with representative (probability) samples, and censuses based on exhaustive data from state data systems.

Estimates from Other Experiments. Previous experiments are an attractive source of estimates of design parameters because they obviously show what can be obtained under realistic conditions with feasible designs (see Schochet, 2005, 2008). On the other hand, previous experiments are limited in serious ways. They represent particular designs (and samples) chosen for particular purposes that may not be the same as those of a new study. There are also relatively few of them that have clearly reported design parameters, making it difficult to find an experiment that closely matches a new study being planned. Obviously, it would be more difficult to find a good match for any study that was studying populations that have not been extensively studied before. Moreover, any particular experiment may not be able to provide information on design parameters such as classroom or district level intraclass correlations (or between-district variance that might be accounted for by blocking by district), although some classroom level intraclass correlation data is provided by Jacob, Zhu, & Bloom (2010).

Even when matching *is* possible, few experiments involve more than a modest number of clusters (e.g., 50 – 80 schools) due to cost considerations. For example, Jacob, Zhu, & Bloom (2010) report design parameters based on an experiment with 23 schools, a demonstration project with 111 schools, and the reading first impact study with 225 schools. The demonstration project was larger than most experiments and the reading first impact study, which was a congressionally mandated study (incidentally, it was not a randomized experiment), is unusually large.

Because the standard error of estimated intraclass correlations and squared multiple correlations (representing variance accounted for at the cluster level) are proportional to the inverse of the square root of the number of clusters, the typical experiment will provide very imprecise estimates of their intraclass correlation or the cluster level variance accounted for by covariates. For example, consider an experiment with a total of 80 schools, 50 students per school, and a true (school level) intraclass correlation of $\rho = 0.20$. The estimate of that intraclass correlation would have an approximate standard error of $SE = 0.028$, so that the two standard error range on either side of the estimate would be plus or minus 0.056[1]. The point here is that *any particular experiment could easily yield an estimate anywhere in that interval*, and the implications for design of an intraclass correlation of 0.14 are quite different from those of an intraclass correlation of 0.26. For example, consider a two-level experiment with no covariates and an effect size of $\delta = 0.30$. A sample size of 29 clusters per treatment group is required to yield a power of 80% if the intraclass correlation is $\rho = 0.14$, but 49 clusters per treatment group are required to yield a power of 80% if the intraclass correlation is $\rho = 0.26$.

The estimate of the variance accounted for at the cluster level in that same experiment would also be imprecise. If the true proportion of variance accounted for were 0.50, the standard error would be about 0.080 and a two standard error interval would be 0.34 to 0.66. The point here is that *any particular experiment could easily yield an estimate anywhere in that interval*, and the implications for design of a covariate that explains 34% of the school level variance are quite different from those of a covariate that explains 66% of the school level variance. For example, consider a two-level experiment with an effect size of $\delta = 0.30$ and an intraclass correlation of $\rho = 0.20$. A sample size of 34 clusters per treatment group is required to yield a power of 80% if the covariate explains 34% of the school level variance, but only 19 clusters per treatment group are required to yield a power of 80% if the covariate explains 66% of the school level variance.

[1] More accurate procedures for computing confidence intervals for intraclass correlations or squared multiple correlations are possible in small samples (see, e. g., Ukoumunne, 2002). We cite this approximation to demonstrate how much estimation error can be expected in estimates from even rather large experiments.

The point here is that even large (a total of 80 clusters) individual experiments cannot provide estimates of design parameters that are precise enough to give meaningful guidance. Moreover, experiments with 80 clusters are larger than many in the literature, and smaller experiments would naturally be even less informative about design parameters.

Estimates from A Few Large Urban School Districts. Some large urban school districts have provided data that has been used to compute design parameters (e.g., Bloom, Richburg-Hayes, and Black, 2007). Such compilations are attractive because they involve well defined samples and larger sample sizes than individual experiments can afford. However they have the weakness that the samples are quite specific. While design parameters so obtained may be reasonable representative of large urban school districts nationally, they are definitely not representative of suburban or rural districts, which includes many of the settings in which education studies are carried out.

Moreover, even moderately large sized districts may not have enough schools to provide highly precise estimates of design parameters. For example in the Bloom, Richburg-Hayes, and Black (2007) study the median number of schools used in any computation was 41, and the largest two districts had 88 and 168 schools (both at the elementary school grade level). Thus the estimates of design parameters they provide are typically no more precise than many large experiments—inadequate, we would argue, to inform precise design decisions.

Also, like the design parameters obtained from experiments, the design parameters obtained from large urban districts have not yet provided information on design parameters such as classroom or district level intraclass correlations.

Estimates Derived from Surveys with Representative (Probability) Samples. Some researchers have used national surveys with probability samples to estimate design parameters (e.g., Hedges and Hedberg, 2007). This strategy has the advantage that the representativeness is well defined, but may be more general (e.g., applying to an entire region or nation) than is desirable for planning some experiments that are more localized. It also typically has the advantage of large numbers of schools, so that estimates can have reasonably high precision.

National survey data generally has the weakness that the areas represented may be more general (e.g., applying to an entire region or nation) than is desirable for planning some experiments that are more localized. Moreover there are few replicated schools within districts, and generally no or weak sampling of classrooms, so they provide no information about district and classroom level intraclass correlations. Finally, many of the survey databases are not recent. Although Hedges and Hedberg (2007) generally found good agreement across surveys of the same population over time, more recent data would obviously be better.

Previous experiments and data from single districts provide estimates of design parameters that are typically too imprecise (and possibly too poorly matched to the current experiment) for detailed guidance. National surveys provide estimates of design parameters that are too general to be optimally useful. Neither provides optimal data on district and classroom level design parameters.

Accomplishments with our Previous IES Grant. The objectives of our previous IES grant were to obtain 5 SLDS datasets, carry out 2, 3 and 4 level analyses of these datasets for each state and specified subsets of the states, and produce a compendium of design parameters that could be made available to researchers. We were able to obtain data from the proposed states and two addition ones for a total of 7 states (Arkansas, Arizona, Florida, Kentucky, Massachusetts, North Carolina, and Wisconsin). We carried out the analyses we specified on all of them.

Journal articles. We published three journal articles:

1. Hedges, Hedberg, and Kuyper (2012) includes a derivation of the large sample variances of the maximum likelihood and restricted maximum likelihood estimators of the intraclass correlation in three and four level models, which we needed to assign sampling uncertainties (standard errors) to the estimates of 3 and 4 level intraclass correlations we were computing.

2. Hedges and Hedberg (2013, in press) provides a basic summary of 3 and 4 level intraclass correlations and multilevel R^2 values for covariates in 7 states. While these are the core results from our analyses, there are many results (e.g., from subsets of schools representing different contexts) that are too extensive to be reported in a single paper or even in a series of papers.

3.

3. Hedberg and Hedges (2013, in press) provides an analysis of heterogeneity of intraclass correlations within states and provides some analyses of associations between district level variables and within-district intraclass correlations.

   Two additional articles are in preparation dealing with heterogeneity parameters and benchmarks for effect sizes associated with year-to-year gains in state assessments that have sufficiently reliable vertical equating to support such analyses.

Conference presentations. We have presented the results of our work at several conferences and workshops including University of Chicago workshops (2013), the STATS-DC 2012 Data Conference, the 2013 Federal Committee on Statistical Methodology (FCSM) Research Conference, the Society for Research on Educational Effectiveness (SREE) 2013 Conference, and a panel at the 2013 Annual Meeting of the American Educational Research Association (AERA). These presentations have been well received and were instrumental in convincing many of the additional states to participate in the study.

Website. We have also developed a website to disseminate detailed results that are too voluminous to be presented in research papers, which has become part of the Northwestern University Institute for Policy Research Q-Center website, assuring it will have a continuing existence irrespective of further grant funding. This tool permits researchers to search for the state, grade, subject matter, etc. of interest and obtain the design parameters that fit their specifications. This website also incorporates information from the previous website that provided design parameters based on national data (that website was called the *Online Variance Almanac*). We note that the *Variance Almanac* has been searched hundreds of times.

Power analysis software. We developed power analysis software that can be used in conjunction with our reference values of design parameters in the form of a Stata macro and released the code for public use (Hedberg 2012b). This program, "rdpower," computes power for a variety of randomized designs: a single level randomized design where there is no clustering, a two-level cluster randomized design where treatment is at level 2, a three-level cluster randomized design where treatment is at level 2, a two-level block randomized design where treatment is at level 1, and a three-level randomized block design where treatment is at level 2. Users can find the program within Stata by typing "findit rdpower."

ICC estimation software. We developed two pieces of ICC estimation software-- "iccvar" that can be used to estimate ICCs and standard errors based on Hedges, Hedberg, and Kyper (2012) and "quickicc" that estiamtes ICCs and standard errors based on Donner and Koval (1980)--and released the code to the public as Stata macros (Hedberg 2012a). The program "iccvar" is a post-estimation command for xtmixed. After fitting a 2-, 3-, or 4-level model with a random intercept (random slopes are not supported), iccvar will calculate the intraclass correlation (ICC) values and the associated standard errors based on the variance components and standard errors of the variance components estimated from xtmixed. The program "quickicc" calculates the intraclass correlation (ICC) after fitting a two-level xtmixed model where the intercept is the only random effect. In addition to calculating the ICC, this program also calculates the standard error of the maximum likelihood large sample ICC (Donner 1980). Users can find the programs within Stata by typing "findit iccvar" or "findit quickicc." As of July 2013, these pieces of software have been downloaded 320 times by various researchers.

IES Training Institute on Randomized Trials. Information from this project and access to the website (and the *Variance Almanac*) have been incorporated into the summer Training Institute on Randomized Field Trials that has been funded by IES for the last 7 years (and was recently funded for an additional 3 years). Participants learn to use these resources to plan randomized trials and are encouraged to use these resources in their teaching and disseminate information about them to other researchers.

Results to Participating States. We have prepared and disseminated results to each of the 7 states that participated, giving them extensive analyses of the design parameters in their states, and explaining how they can be used in planning evaluation studies that involve cluster randomization. For example, the state of Massachusetts has already distributed our estimates to other contractors such as Abt Associates for planning evaluations.

Training for State Research Personnel. We have planned a series of workshops on the use of our software (Stata macros) to carry out the estimation of design parameters, with the hope that these analyses might become routine parts of internal state data analyses. Three states have indicated interest and we anticipate that others will agree to do so.

Technical Assistance to Other Researchers. We have made our Stata macros available to other research teams, such as the one led by Jessaca Spybrook and Joseph Taylor of the Biological

Sciences Curriculum Study (BSCS), to assist with the estimation of ICCs for other academic achievement domains such as science.

In sum, our previous grant helped build capacity in the evaluation community by creating a new database of ICCs for the optimal design of cluster randomized experiments, by providing technical assistance to participating states, and by developing a suite of methodological tools for public use. We anticipate that additional funding will allow us to build on this success..

Why is Another Award Necessary? We believe that additional research is needed for two reasons. First, when we began the previous award, we believed that there would substantial consistency across states in design parameters and that the principle contribution of our research would be to document minor variations and provide more extensive evidence on multilevel design parameters at district, school within district, and classroom within school levels. Our results demonstrated surprising variation in design parameters across states that participated in our project.

Second, we have noted increasing interest in the use of randomized block designs (in part because of resource constraints) that require design parameters involving treatment effects and covariate effects on this variation. As we have noted very little information is available to inform judgments about these design parameters. Moreover treatment effect heterogeneity is increasingly being seen as important for understanding program effects (Weiss, Bloom, and Brock, 2013) and for understanding generalizability of experimental results (see Tipton, 2013).

Estimates from State Longitudinal Data are surprisingly inconsistent. Our previous project proposed to estimate design parameters from state longitudinal data in 5 states. We were actually able to obtain, clean, and analyze data from 7 states (Arkansas, Arizona, Florida, Kentucky, Massachusetts, North Carolina, and Wisconsin), but the results were somewhat surprising. The variation of the intraclass correlations (and R^2 values representing covariate effectiveness) was quite large—large enough to make very substantial differences in the sample sizes required. For example, in grade 4 in reading, the intraclass correlations ranged from 0.013 to 0.127 district level ranged from and from 0.074 to 0.138 at the school-within-district level. The range of state level intraclass correlations at the district level is approximately twice as large as the average across states at any grade (e.g., in reading at grade 3 the mean is 0.049 and the range is 0.94 in grade 3), and the range of school level intraclass correlations is sometimes as large as the average (e.g., at grade 6 in reading the mean is 0.100 and the range is 0.131).

These differences were not the result of statistical estimation error (the standard errors of estimates were typically less than 0.01). Thus the design parameters we estimated for the 7 states we examined are useful for those states, the dataset needs to be extended to additional states to provide a broader set of design parameters for researchers and to better understand the variation between states.

Benchmarks for treatment effect heterogeneity are needed. Treatment effect heterogeneity is crucial for determining the sensitivity of randomized block designs, but it is (by definition) difficult to estimate without actually conducting an experiment because it is specific to the treatment being investigated. Yet researchers must still use *some* values for planning purposes.

What little information is available usually comes from the handful of experiments that happen to have estimated variation in treatment effects of *some* treatment (see e.g., Schochet, 2005, 2008; or Nye, Hedges, and Konstantopoulos, 2000). Such estimates suffer from the same problems of imprecision that plague estimates of intraclass correlations from experiments.

Design parameters estimated in the 7 states in our previous project are useful for those states, but are surprisingly inconsistent. Data from additional states would help us better understand variation among states and would also provide empirical evidence about the heterogeneity of effects across districts, schools and classrooms.

A different approach to this problem is to develop information about heterogeneity of "effects" that interventions might attempt to ameliorate. For example, gaps in achievement associated with race/ethnicity, socioeconomic status, gender, or limited English proficiency are well known. All have all been the target of interventions designed to reduce or eliminate them. Moreover all of these gaps are known to vary somewhat across districts, schools, and classes. While the gaps themselves are not produced by designed interventions, interventions that have been designed to eliminate the gaps are by definition designed to have effects that are as variable as their targets (assuming they could reduce the gaps to zero). While not all interventions are targeted at specific achievement gaps, we believe that establishing a base of empirical evidence about the variation of various gaps would be helpful in grounding estimates of treatment effect variation used in designing studies using randomized block designs. We do not mean to imply that these design parameters are appropriate for research on every intervention, they are in any sense "ideal," or that they should be used mindlessly. However, they are one (and essentially the only) empirical benchmark for grounding thinking about variation of effects in planning evaluation designs. We believe that they represent a better empirical grounding for reasoned argument about the amount of heterogeneity to be expected than wild guesses (or the default values that happen to have been inserted in power analysis software).

Needed: A Broader Collection of Design Parameters Based on State Data Systems. Previous experiments, estimates based on a few schools districts, national surveys, and data from 7 states all have shortcomings in providing precise and relevant estimates of design parameters. The ideal resource would be a comprehensive collection of design parameters based on a census of data on academic achievement in all states. Such censuses exist in many states in the form of state data systems. We have shown the feasibility of exploiting such data systems to develop estimates of design parameters would be based on large samples of schools and classrooms, but be well targeted to the samples of interest to those planning experiments (or quasi-experimental studies that involved clustered sampling designs). Moreover, the state data systems arguably involve the outcome data (state assessments) of greatest policy interest.

We propose to exploit data systems in an additional 7 states (and more states if we can) to compute design parameters (including intraclass correlations, heterogeneity parameters, and variance accounted for by covariates) at the district, school, and classroom level for grade level, and for each dataset where this is feasible. (In some cases the classroom distinction may not be possible or theoretically meaningful, as in high school for reading which is not taught as a separate class.) We will carry out these analyses for entire states and for meaningful subsets of the states defined geographically, in terms of achievement levels (e.g., low achieving schools or districts), and in terms of socioeconomic status (e.g., low SES schools or districts).

Why are District and Classroom Level Design Parameters Necessary? Many researchers carry out design planning (e.g., power analyses) using only two-level designs. One might ask why design parameters involving more than two levels would ever be necessary. We argue that there are at least four reasons as indicated below.

Four level designs (if not analyses) are common. It is important to remember that the actual research design may not always correspond to the planning or analysis (although we would argue that it should). Thus while planning and analysis using two level designs may often be normative practice, a comprehensive review of the designs of experiments funded by IES between 2002 and 2006 revealed that, whatever the planning and analysis used, three and four level designs (typically including some blocking at a level higher than schools) are by far the *most frequent* designs constituting 75% of the total (Spybrook, 2007).

Software can accommodate four level designs. Because designs that involve as many as four levels are common, the software available for power computations, including *Optimal Design* (Raudenbush, Spybrook, Congdon, and Liu, 2006) and *CRT-Power* can handle cluster randomized trials (Borenstein, Rothstein, and Cohen, 2001), both accommodate four level designs. On a practical level, the availability of software that asks for the percentage of variance accounted for by a blocking variable like district will encourage users to provide *some* number, and the number they provide will have a very substantial effect of the power analysis. Some empirical guidance is essential to encourage good scientific practice. An empirical basis for obtaining design parameters is important not only for researchers planning studies but also for reviewers evaluating proposals for funding.

Isn't it OK to skip the classroom level in analysis? In a *completely balanced* three level design, it can be shown that the significance test for treatment effects is valid if the classroom level is omitted (that is, if the data are analyzed as if there were only two levels: schools and individuals). This *does not* mean that the two level power analysis is valid however (at least not with the standard intraclass correlations). Ignoring the classroom level has an impact on the school and individual level variance components so that power computation using the school level intraclass correlation does not give the right answer to the power computation. (Because the variance components are incorrectly defined when a level is omitted, the effective school level intraclass correlation in this analysis depends on details of the design and the class level intraclass correlation in a rather complex way, see, e. g., Moerbeck, 2004.)

Design decisions should be mindful of all sources of variation. We argue that sound experimental design should attempt to be mindful of all sources of variation that affect the experiment. Our experience is that planning a design often involves an iterative procedure whereby several alternatives are considered in terms of cost, feasibility, etc. Concepts like optimal design (optimal allocation of units to levels) can also help inform design decisions. Evaluation of alternatives can best be done when there is empirically grounded information about parameters that impact power, precision, and efficiency of candidate designs.

Research Plan

Primary Research Questions. We propose to use universe-level state data to provide tables of design parameters (including intraclass correlations, effect heterogeneity, and proportions of variance explained by covariates at the district, school, classroom, and individual levels of analysis), for entire grades within states and in a variety of subsets of states that may be relevant to the needs of researchers and of state and local education agencies. Our project primarily seeks to understand variance structures of achievement across many different dimensions. By variance structure we mean the decomposition of the total variation of a particular outcome or effect into district, school, classroom, and individual (student) level, variance components. These variance components can then be used to estimate intra-class correlations and heterogeneity parameters, which are necessary elements in power, precision, and efficiency calculations for cluster randomized trials. Our study seeks to better understand variance structures for students in particular grades in each state subdivided by three dimensions.

The impact of covariates. The first dimension is the set of covariates used in the evaluation designs. Researchers employ a variety of designs in their evaluations, but we will focus on four designs defined by the covariates used. These are designs using

- No covariates (unconditional models)
- Demographic covariates (e.g., race, ethnicity, SES, LEP status, etc.)
- Pretest covariates (models that include previous test scores as covariates)
- Combination models that include both pretest and demographic covariates

Wherever possible, we will evaluate design parameters for all four of these design types.

School context. The second dimension is the school context. Research often evaluates interventions in particular school contexts, such as low-performing schools, impoverished schools, or schools with large minority populations. Sometimes this is because the contexts are of interest in themselves, in other cases it is because interventions, products, and services are targeted at those particular school contexts. Whatever the reasons, the design parameters are often quite different in these context than in the general population of schools, for example national school level intraclass correlations are much lower in low achieving schools than in all schools (see Hedges and Hedberg, 2007). Consequently, we will also calculate variance components for these subpopulations for design planning in each type of school context.

Thus, our first research question for our primary objective is: What is the set of grade and subject specific variance components, and intraclass correlations, for cross-sectional achievement at the district, school, classroom, and student, levels? Our second research question is: What is the set of grade and subject specific variance components, and intraclass correlations, for achievement gains at the district, school, classroom, and student, levels? Our third research question is: How do variance components, and intraclass correlations, differ for each design models? Our forth research question is how do the estimated variance components, and intraclass correlations, differ across different school context?

Secondary Research Questions. We propose to use universe-level state data to provide tables of heterogeneity parameters to empirically ground choices of heterogeneity parameters for randomized block designs (including effect heterogeneity, and proportions of variance in heterogeneity explained by covariates at the district, school, classroom, and individual levels of analysis), for entire grades within states and in a variety of subsets of states that may be relevant to the needs of researchers and of state and local education agencies. We will examine heterogeneity of gender effects, race/ethnicity effects, free or reduced price lunch (FRL) as a proxy for socioeconomic status effects, and limited English proficiency effects, each of which will provide a separate benchmark for empirical evidence about heterogeneity of effects. The details of the analyses will parallel those for the primary research questions.

Data Availability. The previous grant started with five states: Arizona, Florida, North Carolina, Kentucky, and Massachusetts. During the course of the grant, we gave presentations that led two additional states, Arkansas and Wisconsin, to provide data. Over the summer we signed data agreements with three additional states (Colorado, Kansas, Louisiana; see Appendix C) and secured informal agreements to obtain data from five states (Washington, Pennsylvania, Nebraska, Ohio, and West Virginia). Thus, we expect to have secured longitudinal student-level data from eight additional states by the time of award. This will yield eight new sets of state ICC estimates and a total of 15 sets of heterogeneity estimates. Given our success in obtaining cooperation from such a large number of states in our prior work, we anticipate little difficulty in obtaining the data required for this new project to be successful if funded.

Data management. Project staff, particularly Dr. Hedberg, have extensive experience working with state education agencies and their administrative data from NORC's evaluation of the Growth Model Pilot Project (GMPP) under No Child Left Behind[2], our previous ICC work described above, and Hedberg's work with the NCES Education Statistics Support Institute Network (ESSIN) on the creation of a SLDS research agenda. We have requested data elements similar to those used for the GMPP evaluation, including student test scores, school, and district identifiers. For states that have systems in place to link student and teacher data, we will collect teacher identifiers as well. We also will request student-level membership indicators for the following demographic groups: American Indian, Asian, Black, Hispanic, White, Disabled, Limited English Proficiency, and Free-or-Reduced Priced Lunch (as our indicator of SES). NCES school and district identifiers will be requested as well so we can link these data to extant sources such as the Common Core of Data (CCD).

As an example, we requested that North Carolina supply us with elements from three of their state systems.[3] The first is their accountability history system (or ACC_HIST), which is a long file that contains all the reading and mathematics test scores of the current cohort of students in the North Carolina public school system. Data elements from this file will be linked to another system (or ACC_Demo) that houses indicators of student membership in the various demographic subgroups. Finally, teacher identifiers will be gathered from a third membership data system (or ACC_Mem) and linked to the student data. This data set will then be linked to other extant data sources such as the CCD in order to collect additional school-level information such as geographic local and school size. Using the CCD will allow us to both gauge the quality

[2] See http://www2.ed.gov/rschstat/eval/disadv/growth-model-pilot/index.html

[3] See the links at https://www.rep.dpi.state.nc.us/adb/prod2008/tabledocs/ for more information on North Carolina's data systems

of the data we receive from states and gather important school- and district-level control variables without burdening the participating states further.

The CCD allowed us to gauge the quality of state longitudinal data by comparing membership numbers in both data sets after removing charter schools and students with disabilities and de-duplicating SLDS test scores for both reading and math. We considered the state data sound if total membership was within five percent of the CCD estimates. We will continue these checks with future data as it comes in.

Data analysis. Answers to our research questions will require estimation of a number of linear models across several subsamples of data. These analyses will require the use of mixed linear models (also know as Hierarchical Linear Models, or HLM). Mixed models permit estimation of a model in which regression coefficients are taken to vary across groups, and their variation (the variance component for that effect). Our notation for these models will use mixed notation rather than HLM notation because we have so many different models across a number of outcomes that it is simpler to use mixed notation than to introduce another set of equations for different levels of analysis as HLM does.

There will also be variety in the cluster identifiers among the states. Some states, such as North Carolina, will have teacher level identification systems up and running my the time the project receives data, whereas other states such as Washington DC may not have adequate teacher level identification systems. For simplicity, this section will be written assuming optimal data. Where data do not have certain elements, such as teacher identifiers, we will have to adjust our analysis accordingly (e.g., we won't be able to estimate classroom ICCs). We will standardize the test score using the grade-specific mean and standard deviation.

Analyses for Primary Research Questions. The analyses required to address our primary research questions are a series of mixed linear model analyses using variations of the same analysis model. We discuss that model in detail for analyses of achievement status and gains, then discuss the modifications necessary for other analyses necessary to address our primary research questions.

Design parameters for analyses with no covariates. The purpose of these analyses is to establish the set of grade and subject specific variance components, and intraclass correlations, at the district, school, classroom, and student, levels. We describe here the analyses for achievement status (cross-sectional achievement), but the corresponding analyses for achievement gains will be identical expect that they substitute achievement gain for achievement status as the dependent variable. All analyses will be computed for each state separately.

We will fit a mixed model to each state's data that estimates a fixed effect for the average score across each level of analysis and a random effect at each level. Thus, for a k^{th} student in the j^{th} classroom in the i^{th} school in the h^{th} district, the mixed model for score Y_{hijk} for a specific grade and subject will be

$$Y_{hijk} = \mu + \nu_h + \zeta_{hi} + \phi_{hij} + \varepsilon_{hijk}$$

where μ is the average district average, which is the average of each district's school averages, which is the average of each school's classroom averages, which is the average of each classroom's student. The random effect for the h^{th} district is ν_h, the random effect for the i^{th} school in the h^{th} district is ζ_{hi}, the random effect for the j^{th} classroom in the i^{th} school in the h^{th} district is ϕ_{hij}, and the student level residual is ε_{hijk}.

In this analysis, we will estimate four variance components. The first variance component is the district level, or between-district, variance component σ_D^2, which is the variance of ν_h. The second variance component is the school level, or between-school-within-district, variance component σ_S^2, which is the variance of ζ_{hi}. The third variance component is the classroom level, or between-classroom-within-school, variance component σ_C^2, which is the variance of ϕ_{hij}. The final variance component is the individual (student) level, or within classroom, variance component σ_I^2, which is the variance of ε_{hijk}. We will denote σ_T^2 as the total variation, or the sum of all four variance components

$$\sigma_T^2 = \sigma_D^2 + \sigma_S^2 + \sigma_C^2 + \sigma_I^2 .$$

The intraclass correlations at each level of will be equal to the variance component at each level of analysis divided by the total variation. We are aware that there are different possible definitions of the intraclass correlations in three and four level models, but we choose these definitions because they lead to simpler results (given in the previous section) for power, precision and optimal design. Thus, the district level intraclass correlation ρ_D, the school level intraclass correlation ρ_S, and the classroom level intraclass correlation ρ_C are defined as

$$\rho_D = \frac{\sigma_D^2}{\sigma_T^2}, \qquad \rho_S = \frac{\sigma_S^2}{\sigma_T^2}, \qquad \text{and} \qquad \rho_C = \frac{\sigma_C^2}{\sigma_T^2},$$

respectively.

Design parameters for analyses of effect heterogeneity with no covariates. The purpose of these analyses is to establish a set of grade and subject specific variance components for effect heterogeneity, at the district, school, classroom, and student, levels. We describe here the analyses for the effect of gender on achievement status, but the corresponding analyses for other effects will be identical expect that they substitute a different dummy variable or set of dummy variables (in the case of race/ethnicity) for achievement status as the independent variable. All analyses will be computed for each state separately.

We will fit a mixed model to each state's data that estimates a fixed effect for the average score across each level of analysis and a random effect at each level. Thus, for a k^{th} student in the j^{th} classroom in the i^{th} school in the h^{th} district, the mixed model for score Y_{hijk} for a specific grade and subject will be

$$Y_{hijk} = \mu + \nu_h + \zeta_{hi} + \phi_{hij} + (\alpha + \beta_h + \gamma_{hi} + \delta_{hij})T_{ijkh} + \varepsilon_{hijk},$$

where μ is the average district average, which is the average of each district's school averages, which is the average of each school's classroom averages, which is the average of each classroom's student scores. The random effect for the h^{th} district is v_h, the random effect for the i^{th} school in the h^{th} district is ζ_{hi}, the random effect for the j^{th} classroom in the i^{th} school in the h^{th} district is ϕ_{hij}, the random effect for the "treatment" effect in the h^{th} district is β_h, the random effect for the "treatment" effect in the i^{th} school in the h^{th} district is γ_{hi}, the random effect for the "treatment" in the j^{th} classroom in the i^{th} school in the h^{th} district is δ_{hij}, T_{hijk} is the (class mean centered) "treatment" indicator for the the j^{th} classroom in the i^{th} school in the h^{th} district, and the student level residual is ε_{hijk}. The variance component (estimates) for v_h, ζ_{hi}, and ϕ_{hij}, namely σ_D^2, σ_S^2, and σ_C^2 represent the variation of means across districts, schools within districts, and classes within schools, respectively. The variance components for β_h, γ_{hi}, and δ_{hij}, namely σ_{TxD}^2, σ_{TxS}^2, and σ_{TxC}^2 represent the variation of "treatment" effects across districts, schools within districts, and classes within schools, respectively.

The heterogeneity parameters at the district, school, and classroom level, respectively, are

$$\omega_D = \frac{\sigma^2_{T\times D}}{\sigma^2_D}, \quad \omega_S = \frac{\sigma^2_{T\times S}}{\sigma^2_S}, \quad \text{and } \omega_C = \frac{\sigma^2_{T\times C}}{\sigma^2_C}.$$

Design parameters for analyses with covariates. Our second research question is determining the values of design parameters for analyses involving covariates. We envision three kinds of covariates that researcher may employ: pretest only, demographic variables, and a combination of pretest and demographic variables.

Pretest Covariate. We will fit a residualized model to the achievement status (cross sectional) data for each grade and subject. This model will include the previous year's score, X, as a covariate, centered on the state mean at both the student and school level of analysis. For the k^{th} student in the j^{th} classroom in the i^{th} school in the h^{th} district the mixed model for score Y for a specific grade and subject will be:

$$Y_{hijk} = \mu + \beta_1\left(X_{hijk} - \bar{X}_{hij\bullet}\right) + \beta_2\left(X_{hij\bullet} - \bar{X}_{hi\bullet\bullet}\right) + \beta_3\left(X_{hi\bullet\bullet} - \bar{X}_{h\bullet\bullet\bullet}\right) + \beta_4\left(X_{h\bullet\bullet\bullet} - \bar{X}_{\bullet\bullet\bullet\bullet}\right)$$
$$+ v_h + \zeta_{hi} + \phi_{hij} + \varepsilon_{hijk}$$

where, μ is the adjusted district average, $\bar{X}_{hij\bullet}$ is the mean of the pretest scores in the j^{th} classroom in the i^{th} school in the h^{th} district, $\bar{X}_{hi\bullet\bullet}$ is the mean of the pretest scores in the i^{th} school in the h^{th} district, $\bar{X}_{h\bullet\bullet\bullet}$ is the mean of the pretest scores in the h^{th} district, and $\bar{X}_{\bullet\bullet\bullet\bullet}$ is the grand mean of the pretest scores across all schools. The fixed effect β_a is the association between pretest mean at the a^{th} level and the outcome. The (covariate adjusted) random effect for the h^{th} district is v_h, the (covariate adjusted) random effect for the i^{th} school in the h^{th} district is ζ_{hi}, the (covariate adjusted) random effect for the j^{th} classroom in the i^{th} school in the h^{th} district is ϕ_{hij}, and the individual (student) level (covariate adjusted) residual is ε_{hijk}.

In this analysis, we have four estimated variance components. The first is the covariate adjusted between-district variance component σ_{AD}^2, which is the estimated variance of ν_h. The second is the covariate adjusted between-school-within-district variance component σ_{AS}^2, which is the estimated variance of ζ_{hi}. The third is the between-classroom-within-school variance component σ_{AC}^2, which is the estimated variance of ϕ_{hij}. The fourth is the within classroom variance component σ_{AI}^2, which is the estimated variance of ε_{hijk}.

The proportion of variance accounted for by the covariates at each level will also be estimated. Because the analysis of unadjusted outcome yielded unadjusted variance components at each level (σ_D^2, σ_S^2, σ_C^2, σ_I^2), the measure of variance accounted for at the district level R_D^2, the school level R_S^2, the classroom level R_C^2, and he individual level R_I^2 are given by

$$R_D^2 = \frac{\sigma_D^2 - \sigma_{AD}^2}{\sigma_D^2}, \qquad R_S^2 = \frac{\sigma_S^2 - \sigma_{AS}^2}{\sigma_S^2}, \qquad R_C^2 = \frac{\sigma_C^2 - \sigma_{AC}^2}{\sigma_C^2}, \text{ and } \quad R_I^2 = \frac{\sigma_I^2 - \sigma_{AI}^2}{\sigma_I^2},$$

respectively.

Note that it is also possible to compute covariate adjusted intraclass correlations. We will not compute covariate adjusted intraclass correlations, because essentially all the software available for computing statistical power and design sensitivity with covariates requires unadjusted intraclass correlations (and R^2 values). We have found that even sophisticated researchers are sometimes confused about whether to use adjusted or unadjusted intraclass correlations in power or precision computations with covariates.

Demographic Covariates. The demographic will focus on membership in the following five subgroups: Black (BLK), Hispanic (HSP), Disabled (DIS), Limited English proficient (LEP), and eligible for free or reduced price lunch (FRL) as a proxy for socioeconomic status. We may also explore the use of other school level characteristics, such as number of teachers or urbanicity, originating from derived variables (such as the percent of students who are Hispanic) and from the Common Core Data available online. Our analysis of preliminary models will help us decide which school level characteristics are both broad enough to be available to most researchers yet important enough to show a measurable increase in power.

We will fit a residualized model to the achievement status (cross sectional) data for each grade and subject similar to that used when pretest was the covariate, except that this model will include the five demographic variables and their deviations from means at each level of the analysis. Thus there will be 20 fixed effects in the analysis when all of the demographic covariates are included. We plan to run the full analysis with all 20 fixed effects, then explore models with subsets of the covariates to determine which might be the most efficient covariate sets. For example, how much additional variance is accounted for (particularly at the school and district levels, which have the largest impact on power) by adding dummy variables for Black and Hispanic over just using free or reduced price lunch eligibility?

The result of these analyses will be, as in the case of analyses adjusting for pretest as a covariate, will be a set of R^2 values representing variance accounted for at each level (that is R_D^2, R_S^2, R_C^2, and R_I^2).

Pretest and Demographic Covariates. To help determine how much demographic covariates might add to pretest as a covariate, or vice versa, we will carry out analyses using both pretest and demographic variables are covariates. We will fit a residualized model to the achievement status (cross sectional) data for each grade and subject similar to that used when only pretest was the covariate, except that this model will include both pretest and the five demographic variables and all their respective deviations from means at each level of the analysis.

The result of these analyses will be, as in the case of analyses adjusting for pretest as a covariate and demographic covariates alone, will be a set of R^2 values representing variance accounted for at each level (that is R_D^2, R_S^2, R_C^2, and R_I^2).

School Contexts. Our second primary research question is how the design parameters differ across different school contexts. To answer this question we will select subsets of schools from each state. We will then estimates all the models described above on this subset. We will use the following subsets:

- Schools that are classified as low performing by current policy standards (e.g., schools that are not making Adequate Yearly Progress under NCLB)
- Schools with greater than 35% of students who are eligible for free or reduced priced lunch
- Schools with greater than 35% under represented minorities
- Schools with greater than 50% students below state average for reading or math
- Schools in rural areas
- Schools in urban areas.

Estimating design parameters (variance components, ICCs, the effectiveness of covariates at explaining variance at each level) for each of these sub-sets of schools will provide more specific design parameters for studies of specific school contexts and allow researchers to have more precise basis to design studies that will have adequate statistical power and precision.

Why Does This Work Require IES Support? One might wonder why a series of reasonably straightforward analyses requires multi-year research support. Although individual analyses *are* straightforward, the project as a whole involves a great deal of analysis for two reasons. First, the preparation of datasets for analysis always requires significant effort. Second, the combinatorics of our proposed research implies that we will be carrying out thousands (actually tens of thousands) of analyses when subject, grade, state, model, subpopulations are considered. For example, the previous grant produced a final set of 65,039 parameters (*ICCs*, R^2s, variance components, and standard errors), with typically 10,000 parameters per grade across the seven states. These numbers reflect the final run of data and not the numerous prior drafts (our experience is that glitches do occur that require repeating analyses). Thus, at a minimum, thousands of analyses will have to be carried out for each state, and we will have multiple states. Our experience in the past is that this is daunting, but possible.

Scope of Work

We are proposing a 36-month project that will produce reports to IES, academic papers, workshops and technical assistance for participating states. Our schedule matches our successful first grant. The primary task in the first year will be collecting data from participating states and calculating variance components, ICCs, and measures of heterogeneity from this data based on several conditions and for several types of schools and regions. The second year will be an analysis of the results from the first year and the production of reports and guides to help design intervention research. The third and final year will be preparing the material for publication so that state and local education agencies and other researchers are able to use them. We look forward to working with states to better understand their planned or ongoing interventions and to provide technical assistance. Custom software will also be developed to aid users in the design of their studies.

Year 1. The first year will have two main tasks. The first half of the year will be spent collecting data from the participating states. This will involve obtaining permission to collect this data from NORC's Institutional Review Board, meeting the participating state's requirements for data security, and generating the data disclosure and/or memoranda of agreements with participating state education agencies. Given our early success in gaining state participation in this project, we do not foresee any difficulties in obtaining their data or in expanding the list of participating states. Once the data are in hand, we will go through a quality control process of checking the data to make sure the population of students in the state data matches other published sources such as the CCD.

The second half of the year will be the start of the analysis, beginning with the initial draft of the analysis code. We will use Stata 13's MIXED procedure to estimate our variance components and Hedberg's (2012) ICCVAR software, developed for the first grant, to estimate the ICCs and standard errors. Stata is an ideal choice because of its ability to keep estimated statistics, such as variance components, in memory and manipulate them. This means that all the variance components, ICCs, R^2 values, and effect size tables will be produced automatically, eliminating the chance of human error in table production. This code will form the basis of a Stata program that will be provided to participating states so that they are able to calculate these statistics for their own use. A more detailed breakdown of activities in the first year follows:

Months 1-3. The first months will be primarily spent fulfilling required administrative tasks such as obtaining IRB approval, setting up NORC resources for housing and storing the data, and gathering the data from participating states. Each state has their own security and data extraction protocols, so much of time will be spent working with state officials to gather data. This requires that NORC enter into Memoranda of Agreements (MOAs) with the participating states clearly specifying the terms of using the data and the project deliverables. Once the MOAs are in place, we will work with states to select which variables, or tables, we need extracted from their system. As described in our data section, these will include test scores, demographic information, and identifiers required for district, school, and classroom levels of analysis. The states will then perform the extracts by the end of the third month.

Months 4-6. Once the data have been extracted, we expect the states will load their data into NORC's SFTP (secure file transport protocol) servers or send encrypted disks. Data received will be stored on NORC's secure network servers and any hard copies will be destroyed. It will then be transferred into Stata format for cleaning and processing. This will include checks for missing data and for discrepancies between state-supplied data and other sources such as CCD concerning the numbers of districts, schools, and students in each subgroup and for each grade. Any discrepancies found will be referred back to states for resolution. We also will send letters to states that decided not to participate. These letters will inform them that the project was funded and invite them to participate again. Should any of these states decide to participate, we will follow the same process to secure access to the data.

Months 7-9. After the data are in hand and checked for quality, principle programming will begin. The first step in the programming process will be any cleaning of non-applicable test scores, such as special tests for students with disabilities. We will also write programs that standardize the data from each participating state into a common set of variables and membership codes. At this time, the data will be linked to any extant sources such as the CCD. Again, quality checks will be performed to make sure data were not censored due to the cleaning and standardization process.

Months 10-12. At this point the data received from the states will have been cleaned and standardized, thus the focus of this time period will be running the unconditional models. This will allow us to build our initial technology for creating tables of variance components, intraclass correlations, and effect sizes. All estimations and tables will be produced Excel-ready in Stata.

Year 2. The second year will be focused on the analysis and writing up the preliminary results. It also will include a second round of data collection from participating states. Based on our experience working with national samples for the Hedges and Hedberg (2007, 2013) papers, we expect that calculating the statistics of interest and running quality checks will require considerable time and effort. While we already have code in place for the ICC estimates, we will need to generate new code for the heterogeneity estimates. Thus we expect the code writing process to extend into the second year as we finalize our Stata programs but that the code will be finalized and useable results will be obtained by the end of the year.

The other primary task of the second year will be distribution of the preliminary results, including draft reports to IES and presentations at professional meetings. The first report to IES will showcase tables of variance components, intraclass correlations, R^2 values, and heterogeneity estimates, from each participating state. A second report will showcase tables of the effect sizes for year-to-year gains for states with applicable scales. Both reports will be provided to participating states as well. We also plan to write two papers that will mirror the results of the reports to IES and additional papers outlining results for each analytical context across states for presentation to academic audiences. A more detailed breakdown of activities in the second year follows:

Months 13-15. These months will be spent two tasks. First, we will extend our code to include the other models that involve covariates. The code for these models will be somewhat different by including covariates and requiring calculation of R^2 values. Our previous experience shows that calculating the R^2 values requires an additional set of unconditional models that use the exact same case base as the adjusted models. Second, we will quality check the code that produces the heterogeneity estimates. Since this was not done in the previous study, we expect to spend a substantial amount of time ensuring that this code is of the same quality as our ICC code.

Months 16-18. With all of the base code written at this point, the task for these months will be quality checking the final code and finalizing the format of the estimation tables.

Months 19-21. These months will be focused on submitting abstracts for presentations of our initial findings at professional association meetings.

Months 22-24. This time will be spent on two main tasks. First, we will write and submit a draft report to IES presenting the results of our analyses. Second, we will take our core code and begin the process of compiling a general piece of Stata software to calculate the heterogeneity statistics using a specifically formatted dataset. This will be the first piece of Stata software developed for this project and will be delivered as part of our technical assistance to states so that they can compile their own tables of variance components, intraclass correlations, R^2 values, effect sizes, and heterogeneity measures.

Year 3. The third year will primarily focus on deliverables. We envision five deliverables for the third year: final reports to IES, power analysis software, a workshop for participating states, academic papers of results, and a website available to IES and the research community. This website will expanded version of the website developed under the previous grant and hosted at the Northwestern University Q-Center. A more detailed breakdown of activities in the third year follows:

Months 25-27. We will test our first stand-alone program for compiling variance components, intraclass correlations, R^2 values, effect sizes, and heterogeneity measures, on any new state data we receive. This will allow us to finalize our software. We also will meet with our Web developers during this period to begin disseminating our findings though an interactive Web portal. This website will allow users to select states and other parameters, such as research contexts, and receive tables of effect sizes and required sample sizes in return. The scheduling of state workshops and technical assistance sessions also will begin.

Months 28-30. This period will be focused on adjusting the multi-level power analysis software to incorporate new functions such as minimum detectable effect sizes and using heterogeneity measures. We also will follow up with state officials to develop agendas for the workshops and to coordinate any technical assistance they may request. Draft reports and academic papers will be finalized as well.

Months 31-33. The power analysis software will be completed during this period, as will the dissemination website and the final tables of results. We also will host the state workshop and will provide any technical assistance requested by states.

Months 34-36. These final months will see the delivery of three papers. The first paper will outline the set of variance components, intraclass correlations, and R^2 values for all participating states. This paper will use the Hedges and Hedberg (2007, 2013) pieces as a reference point to discuss any differences found. The second paper will be a follow up to the Hedberg and Hedges piece that investigated the association between ICCs and district structure. Finally, we will use our large sample of states to write a paper investigating the differences across states and with our national estimates.

Personnel

The team at NORC is especially suited for this project because of their expertise in this subject matter and their experience working with state administrative data.

Larry Hedges, Principle Investigator. Dr. Hedges is a national leader in the fields of educational statistics and evaluation. He joined the Northwestern faculty in 2005 where he is one of eight Board of Trustees Professors, the university's most distinguished academic position. He holds appointments in statistics (where he is chairman elect), psychology, and education and social policy. Hedges will contribute 2 months per year to this project.

Hedges is best known for his work to develop statistical methods for meta-analysis, but has extensive experience with large scale data analysis and randomized trials. For the last 7 years he has been director of the IES funded Summer Training Institute in Randomized Field Trials for Established Researchers. Widely published, he has authored or co-authored approximately 200 journal articles and nine books, including the seminal Statistical Methods for Meta-Analysis (with I. Olkin) and The Handbook of Research Synthesis (with H. Cooper). Hedges is a member of the National Education Sciences Board. He is an elected member of the National Academy of Education and is a fellow of the American Academy of Arts and Sciences, the American Statistical Association, the American Psychological Association, and the American Educational Research Association. He is president of the Society for Research on Educational Effectiveness, and past president of the Society for Research Synthesis Methods. He served as Editor of the Journal of Educational and Behavioral Statistics. Hedges was elected Statistician of the Year by the Chicago Chapter of the American Statistical Association in 2013/14. He is also a NORC Senior Fellow.

Eric Hedberg, Co-Principle Investigator. Dr. Hedberg is a Senior Research Scientist at NORC who has worked on a number of large-scale research projects in education involving analysis of raw data from a number of state education agencies, federal data, survey data such as NELS-88, National Longitudinal Survey of Youth, ECLS, Trends in International Mathematics and Science Study, and the National Assessment of Educational Progress state and federal data. Hedberg has been a successful Co-PI of the first ICC grant project and has been instrumental in a number of recent NORC education contracts such as the analysis of the new ECLS-K 2011 data in the report "Change and Stability at the Starting Gate: A Comparison of America's Kindergartners in 2010 and 1998." He has experience and expertise both in managing data, and designing rigorous scientific studies of education data. He has several of years of experience conditioning raw administrative data into research databases appropriate for policy analysis, and performing such analyses.

Dr. Hedberg specializes in generalized linear models and mixed regression modeling. He has co-authored papers on experimental design and presented at several professional conferences, including methodological meetings. He has been a source of technical assistance for IERI projects through the Data Research and Development Center. He has served as methodologist for a number of successful research proposals, built sampling frames and calculated weights. He has worked as task leader for analysis on the GMPP evaluation project, estimating the effect of growth models on AYP determinations for schools in nine states. Hedberg has also authored several Stata estimation commands with applications ranging from meta-analysis to complex sample variance estimation such as variance estimation for percentiles using replicate weights. He has co-authored several methodological pieces that have appeared in education, medical, and criminological journals. Currently, in addition to the IES funded research on ICCs, he serves as a methodologist for the Minnesota Reading Core Evaluation, the Survey of Doctorate Recipients (calculating several detailed statistical tables that will be published by the NSF), and the Evaluation of the Graduate Research Fellowship Program (designing the quantitative evaluation and propensity models).

Resources

NORC maintains a production-ready network of servers designed to reliably support all NORC activities. These include file, application, and database servers, as well as Web hosting, and a Storage Area Network (SAN) providing expandable data storage. The NORC Wide Area Network (WAN) links PC workstations attached to Local Area Networks (LANs) across all NORC offices, providing users with access to more than sufficient online disk storage, as well as the various file, print, and application servers; UNIX hosts; and high-volume laser printers. NORC's data center is supported by an uninterruptible power supply (UPS), resulting in zero downtime in the event of power outages. In addition, the Virtual Private Network (VPN) structure links staff from remote locations throughout the United States.

NORC's offices have voicemail, teleconferencing, and videoconferencing installations in order to facilitate formal and informal communications. All NORC office staff use high-performance, multimedia desktop PCs and the Microsoft Windows operating system. Facilities are interconnected with high-speed telecommunications for voice and data transmission. NORC remains vigilant in frequently upgrading and expanding equipment as more advanced technologies become available.

Security is a major concern when working with educational data. NORC and the research team have experience in the proper handling and storage of sensitive data. We will de-identify the data, store raw data in secure rooms with locked cabinets, use only secure servers for analysis of cleaned data, and produce only tables of aggregated statistics. Not only will be follow NORC security procedures but we will also conform to the security requirements of participating states.

NORC's security program is compliant with federal government regulations and can be adapted easily to meet the unique requirements of any project. Recent years have been highlighted by breaches in computer security of various government agencies as well as private industry. As a result, every organization has been challenged to meet the potential for such security issues with multilayered approaches to securing computer systems and the data they contain. NORC takes the matter of computer security seriously and has developed a multi-tiered approach to managing the issues surrounding computer and data security.

On many of our projects, compliance with NIST 800.53 recommendations is a requirement. NORC currently has other government projects that require similar compliance, and recent audits by those projects have found that our systems meet or exceed these requirements. We have projects underway for the Department of Labor, the Federal Reserve, and the Bureau of the Census that require independent audits to confirm compliance. In each case we have successfully met the NIST standards.

NORC takes great care to enforce physical security measures specifically designed to ensure that access to confidential data is restricted to only those employees who possess the need, as well as the authorization, to review such information. NORC requires the use of internal network data storage services to store all project-related datafiles. Partitioned network storage is provided for each project to mitigate the potential for data loss due to accidents, computer equipment malfunction, or human error, as well as to administer access rights regarding privacy issues related to both legal and contractual obligations. Wide arrays of network security precautions are undertaken by NORC to ensure the proper storage of all project data.

All remote access to internal NORC computing resources requires two-factor authentication and encrypted channels. Only secure, encrypted file transfers are used when exchanging files with clients and/or partners over the Internet. All of NORC's laptop computers are provisioned with an automatic full disk encryption system to protect against loss of sensitive data should any of these machines be lost or stolen. All user credentials and associated access permissions are subject to the controls and standards maintained by NORC's IT department. In particular, passwords must meet stringent requirements for length and complexity and must be changed on a regular basis.

References

Bloom, H. S. (1995). Minimum detectable effects: A simple way to report statistical power of experimental designs. Evaluation Review, 19, 547-556.

Bloom, H.S., Richburg-Hayes, L., & Black, A.R. (2007). Using covariates to improve precision: Empirical guidelines for studies that randomize schools to measure the impacts of educational interventions. Educational Evaluation and Policy Analysis,29: 30-59.

Borenstein, M., Rothstein, H., & Cohen, J. (2001). Power and Precision. Teaneck, NJ: Biostat.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences (2nd ed.). New York: Academic Press.

Donner, A. & J. J. Koval. 1980. The Large Sample Variance of an Intraclass Correlation. Biometrika. 67:3. 719-722.

Donner, A., & Klar, N. (2000). Design and analysis of cluster randomization trials in health research. London: Arnold.

Hedberg, E. C. (2012a). ICCVAR: Stata module to calculate intraclass correlation (ICC) after xtmixed. Statistical Software Components.

Hedberg, E. C. (2012b). RDPOWER: Stata module to perform power calculations for random designs. Statistical Software Components.

Hedberg, E.C., Hedges, L.V. & (2013). Empirical Estimation and Prediction of Within-district Intraclass Correlations of Academic Achievement in Elementary Grades. Evaluation Review. In Press.

Hedges, L. V. (2007). Effect sizes in cluster randomized designs. Journal of Educational and Behavioral Statistics, 32, 341-370.

Hedges, L. V., & Rhoads, C. (2010). Statistical power analysis in education research. Washington, DC: US Department of Education.

Hedges, L. V., Hedberg, E. C., & Kuyper, A. M. (2012). The Variance of Intraclass Correlations in Three-and Four-Level Models. Educational and Psychological Measurement, 72(6), 893-909.

Hedges, L.V., & Hedberg, E.C. (2007). Intraclass correlations for planning group- randomized experiments in education. Educational Evaluation and Policy Analysis, 29, 60-87.

Hedges, L.V., & Hedberg, E.C. (2013).Intraclass Correlations and Covariate Outcome Correlations for Planning 2 and 3 Level Cluster Randomized Experiments in Education. Evaluation Review. In Press.

Hedges, L.V., & Rhoads, C. (2009). Statistical Power Analysis in Education Research. U.S. Department of Education. Washington, DC: National Center for Special Education Research, Institute of Education Sciences.

Holland, P. W., & Rubin, D. B. (1983). On Lord's paradox. In Wainer & Messick (Eds.), Principals of modem psychological measurement: A Festschrift for Frederic M. Lord (pp. 3-25). Hillsdale, NJ: Erlbaum.

Jacob, R., Zhu, P., & Bloom, H. (2010). New empirical evidence for the design of group randomized trials in education. Journal of Research on Educational Effectiveness,3, 157-198.

Konstantopoulos, S. (2008a). The power of the test for treatment effects in three-level block randomized designs. Journal of Research on Educational Effectiveness, 1,265-288.

Konstantopoulos, S. (2008b). The power of the test for treatment effects in three-level cluster randomized designs. Journal of Research on Educational Effectiveness, 1,66-88.

Konstantopoulos, S. (2009). Optimal design in three level cluster randomized experiments. Evaluation Review, 33, 335-357.

Moerbeck, M. (2004). The consequences of ignoring a level of nesting in multilevel analysis. Multivariate Behavioral Research, 39, 129-149.

Murray, D.M. (1998). Design and analysis of group-randomized trials. New York: Oxford University Press.

Nye, B., Hedges, L. V., & Konstantopoulos, S. (2000). The effects of small classes on academic achievement: The results of the Tennessee class size experiment. American Educational Research Journal, 37(1), 123-151.

Raudenbush, S. W., Martinez, A., & Spybrook, J. (2007). Strategies for improving precision in group-randomized experiments. Educational Evaluation and Policy Analysis, 29(1), 5-29.

Raudenbush, S., Spybrook, J., Congdon, R., & Liu, X. (2006). Optimal Design for Longitudinal and Multilevel Research: Documentation for the "Optimal Design" Software. William T. Grant Foundation.

Raudenbush, S.W. (1997). Statistical analysis and optimal design for cluster randomized experiments. Psychological Methods, 2, 173-185.

Raudenbush, S.W., & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. Psychological Methods, 5, 199-213.

Rhoads, C. (2011). Implications of "contamination" for experimental design in education. Journal of Educational and Behavioral Statistics, 36, 76-104.

Schochet, P. (2005). Statistical Power for Random Assignment Evaluations of Education Programs. U.S. Department of Education. Washington, DC: Institute of Education Sciences.

Schochet, P. (2008). Statistical power for random assignment evaluations of education programs. Journal of Educational and Behavior Statistics,33, 62-87.

Spybrook, J. (2007). Experimental designs and statistical power of group randomized trials funded by the Institute of Education sciences. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Spybrook, J., & Raudenbush, S. W. (2009). An examination of the precision and technical accuracy of the first wave of group-randomized trials funded by the Institute of Education Sciences. Educational Evaluation and Policy Analysis, 31(3), 298-318.

Tipton, E. (2013). Improving Generalizations From Experiments Using Propensity Score Subclassification Assumptions, Properties, and Contexts. Journal of Educational and Behavioral Statistics, 38(3), 239-266.

Ukoumunne, O. C. (2002). A comparison of confidence interval methods for the intraclass correlation coefficient in cluster randomized trials. Statistics in Medicine, 21, 3757-3774.

Weiss, M. J., Bloom, H. S., & Brock, T. (2013). A Conceptual Framework for Studying the Sources of Variation in Program Effects.