

Support for Knowledge Accumulation in Educational Research

Barbara Schneider

Factors that Bear on Knowledge Accumulation in Educational Research

- The field of education has become increasingly diverse and fragmented.
- There are different understandings of what constitutes scientific inquiry, evidence, and interpretation

Promoting Knowledge Accumulation in Educational Research

- Replication presents an important unifying concept and a starting point for constructive dialogue.
- Data sharing is essential to replication and is a vehicle through which investigators can build upon designs, create and revise measures, and study different populations for purposes of developing new theories.

The Importance of Replication

Without convergence of results from multiple studies, the objectivity, neutrality, and generalizability of research is questionable.

Mechanisms that Facilitate Replication

- Upholding ethical standards for replication and data sharing in the professional associations
- Reinforcing these standards by requiring researchers to abide by them when publishing findings in professional journals
- Maintaining institutional infrastructures that assist researchers in data sharing and provide opportunities for study replication

Professional Ethical Standards

ASA Ethical Standards

- Sociologists make their data available after completion of the project or its major publications, except when proprietary agreements with employers, contractors or clients preclude such accessibility or when it is impossible to share data and protect confidentiality.

ASA Ethical Standards

- Sociologists anticipate data sharing as an integral part of a research plan whenever data sharing is feasible.

ASA Ethical Standards

- Sociologists share data in a form that is consonant with the research participants' interests and protect the confidentiality of the information they have been given. They maintain the confidentiality of data, whether legally required or not; remove personal identifiers before data is shared; and if necessary use other disclosure avoidance techniques.

ASA Ethical Standards

- Sociologists who do not otherwise place data in public archives keep data available and retain documentation relating to the research for a reasonable period of time after publication or dissemination of results.

ASA Ethical Standards

- Sociologists may ask persons who request their data for further analysis to bear the associated incremental costs, if necessary.

ASA Ethical Standards

- Sociologists who use data from others for further analyses explicitly acknowledge the contributions of the initial researchers.

(ASA Code of Ethics, 1997)

Data Sharing in the Journals

Reinforcing Norms for Replication in Journals

- In the physical and social sciences, norms for replication are reinforced in professional and disciplinary journals.
- Authors are often required to follow well-documented and rigorous procedures for data sharing when a paper is accepted for publication

Data Sharing in Journals

- *Science* and *Nature* both require that data be arrayed in identified files that directly correspond to results reported in the tables and figures in the manuscript.
- *Nature* requires that any supporting data set for which there is no public repository be made available to any interested reader on and after the publication date.

Data Sharing in Journals

- *Demography* requires that authors preserve the data used in their analyses and make that data available to others at a reasonable cost for three years after publication.

Building an Infrastructure for Data Sharing

National Center for Education Statistics

- The government has several different Centers and Institutes designed to make data accessible to researchers, such as the National Center for Education Statistics (NCES).

Inter-university Consortium for Political and Social Research

- The Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan maintains and provides access to a vast archive of social science data for research and instruction, and offers training in quantitative methods to facilitate more effective use of these data.

The Murray Center

- The Murray Research Center at the Radcliffe Institute for Advanced Study at Harvard University includes both survey and qualitative data in a national archive of materials from over 270 studies.
- The Murray Center focuses on social science data, especially qualitative and longitudinal data on women's lives, human development, and the life span.
- The Murray Center includes materials like case histories, open-ended interviews, and response to projective tests.

Incentives for Data Sharing

- There has been some discussion about requiring recipients of research grants to provide a public use data file at the completion of the work.
- This type of incentive for data sharing could be very useful for ensuring norms of replication.

Replication in Education

Quantitative Examples

- The release of the Coleman Report was met with an onslaught of criticism.
- A year-long faculty seminar was organized at Harvard University to review the findings and reanalyze the data.

Quantitative Examples

- When the analyses of the High School and Beyond data became public, the findings were again questioned.
- Reanalysis took place in many universities, most notably at Stanford University.

Quantitative Examples

- Two other studies, both controversial, have been subject to reanalysis: the Tennessee STAR Experiment and the New York Voucher Experiment. In neither case have the data been widely available to researchers.

Replication often occurs because
of controversy.

Replication Using Qualitative Data

Qualitative Examples

- The Three-City Ethnography, funded by NICHD, is designed to conduct fine-grained assessments of how, over-time, welfare reform policies influence the day-to-day lives of low income African American, Latino, Hispanic, and non-Hispanic white families.
- The ethnography team consists of over 210 research scientists, ethnographers, qualitative data analysts, system programmers, and staff.

Qualitative Examples

- The Three-City Ethnography has generated an extensive qualitative data set consisting of over 45,000 pages of field notes and supporting data (tapes, diagrams), which have been organized into a consistent data management system.
- Data transfer, tracking, and storage mechanisms are in place that protect the data from corruption and loss and maintain respondent confidentiality.

It would appear that strong training in collaborative research -- where the work of investigators is routinely scrutinized and graduate students are fully engaged in the research, as in the instance of the Three-City project -- helps to minimize personal risk while maximizing the opportunity for replication and strengthening the discipline.

Facilitating Data Sharing

- The large-scale data sets that are based on probability samples, such as ECLS, NELS:88-92, and Baccalaureate and Beyond, are potentially useful for testing and constructing measures, contextualizing the results of smaller-scale studies, and training graduate students.

Problems with Large-scale Data Sets

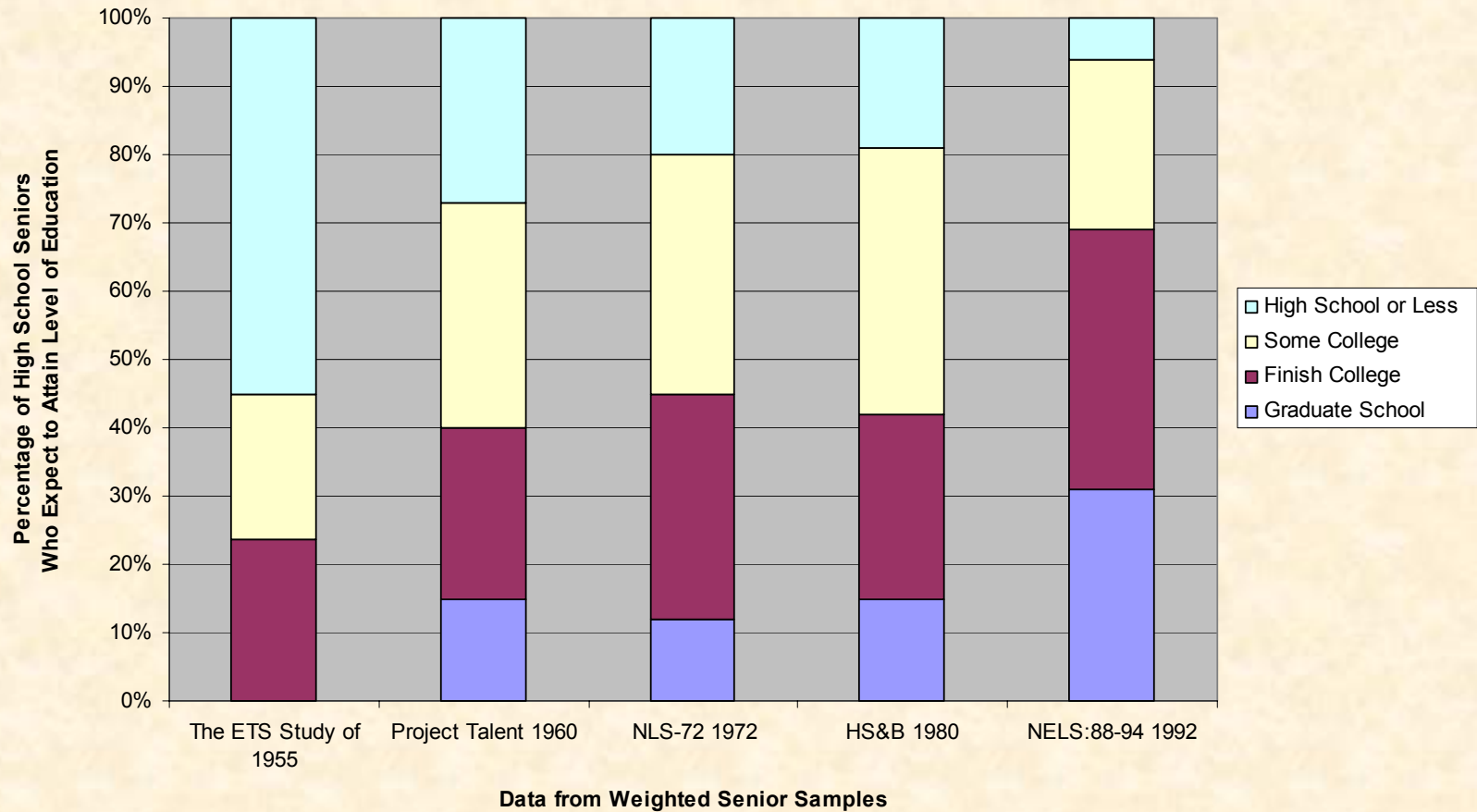
- Investigators do not incorporate advances in the field when developing new questions.
- Investigators are often uninformed about statistical advances that improve the analytic capacity of the datasets.

Advances in Analytic Techniques

- Procedures have been developed for addressing problems of selection bias, and imputation techniques allow investigators to avoid problems associated with missing data.
- Advances in statistical techniques are increasing the accuracy and power of multivariate analyses using longitudinal data.
- Probability samples with extensive item pools are one of the best indicators of what our population looks like.

Example

The Ambitious Generation: America's Teenagers, Motivated but Directionless



- Large-scale probability samples can be combined to create new measures that benefit the discipline.
- National data sets can be used to contextualize smaller-scale data sets.
- Large-scale data sets provide excellent opportunities for training young scholars in quantitative analyses. They are especially helpful for teaching statistical methods.

Interagency Education Research Initiative

- There are approximately 60 funded IERI research projects focused on educational advancements and intervention effectiveness.
- Using a variety of research designs that often use random assignment of subjects, observation, surveys, and other methods, these studies are producing new information on learning, instruction, and achievement.
- IERI projects on a the whole are innovative, systematically designed and implemented interventions that are required to produce scientific evidence on US classrooms.

Data Research and Development Center

- The Data Research Development Center (DRDC) is committed to improving the research capacity of faculty, students, and scholars engaged in interdisciplinary work in the areas of learning, instruction and achievement.
- The DRDC is working with other IERI researchers to determine what programs and interventions are most effective for different students, using IERI and other national data sets.

- Over the past year DRDC has worked to build a community of scholars who share designs, measures, and results.
- We have tried to learn how individuals interact with each other and who they rely on for advice in their research, in training students, and in their professional lives.
- We anticipate learning what venues strengthen information-sharing projects.

- DRDC has started to build common measures that researchers can use across data sets. We are currently examining how factors such as socioeconomic status are measured across data sets.
- We are trying to promote analysis of multiple data sets.

DRDC is providing several kinds of technical assistance to investigators, including:

Study Design

Sampling expertise: helping investigators to replicate results of promising interventions using larger samples

Access, cooperation, and assignment: providing information on conducting scientific research in schools

Instrumentation

- Assessing fidelity of implementation: sharing designs, instruments, and items for measuring what conditions affect the implementation of the intervention and its potential for replication in other contexts.

Analysis

- Analytic solutions: advising on what analytic procedures, such as hierarchical linear analysis, would be the most useful for measuring the effectiveness of the intervention

Building a Professional Community

- To further the development of the professional community, we have created a website and are working with the computer science department at the University of Chicago to establish virtual collaborations whereby investigators will be able to interact with each other through learning access grids and other technologies.

Looking to the future

- We hope that the activities of DRDC, both the successes and the failures, will make a difference in creating a stronger more open scholarly community in the field of education.
- Our main goal is to strengthen research norms in education that encourage the systematic study of education.