Spring 2014 Volume XXXIV, No. 2

ICPSR BULLETIN

IN THIS ISSUE

Summer Program offers new workshops 4
Interest in Measures of Effective Teaching data grows 6
New ICPSR Council members have first meeting together 10
Data from '50s Cornell Study of Occupational Retirement available 11

New public-access research data sharing service launched

openICPSR provides means for data depositors to fulfill public-access obligations of grant or contract RFPs

By Mark Thompson-Kolar
ICPSR Editor

As the scientific community's focus on public access to data has grown, ICPSR in February launched its research data sharing service, openICPSR. It allows the public to search for and access public-use social and behavioral science research data at no charge.

The service is available in a beta version through July 2014. During this period, member institutions can deposit data into it without a charge. The full launch, which will include a system for payment at deposit, will take place when the beta period concludes.

"openICPSR is a new way of sharing data for ICPSR," said ICPSR Director George Alter. "It's largely a response to changes in Washington and funders, where they're telling researchers, 'Government-funded data should be public.' "

In the past three years, federal agencies have introduced requirements for public access to federally funded research data in two key documents. In January 2011, the National Science Foundation outlined a requirement that all proposals for funding must include a data management plan; the National Institutes of Health has a similar requirement. In February 2013, the White House's Office of Science and Technology Policy issued a memorandum directing federal funding agencies with an annual R&D budget above $100 million to develop plans to support increased public access to the results of federally funded research.

openICPSR is available to any researchers who want to make their

Open */ Continued on page 2*

Open */ Continued from page 1*

data publicly available in a sustainable archive, and it provides a means for data depositors to fulfill the public-access obligations of grants or contracts.

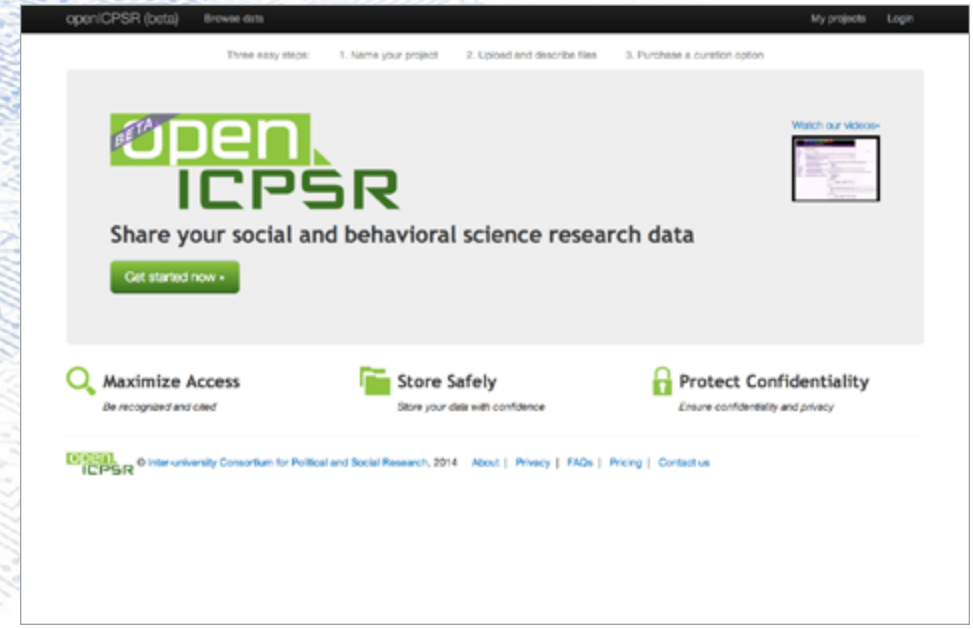
"We've created a new mechanism so that researchers can meet federal requirements and be confident that their data will be visible in the ICPSR catalog and preserved for the future," Alter said.

The service is entering into an information space that recently has attracted other public-access data sharing services. "We believe that researchers want to put their data in a repository that understands social science research. ICPSR has been working with the social science community for more than fifty years, and we know that a dataset cannot be documented in the same way as a book or a video."

"As a global leader in data stewardship, it was important for us to adapt to the changing data sharing environment," said Linda Detterman, Marketing and Membership director. "openICPSR isn't just a catcher's mitt for data. We didn't want to just throw something out there. We wanted to make sure researchers' data were well taken care of and findable."

openICPSR offers three deposit options.

• The $600 Self-Deposit Package enables research scientists to deposit data rapidly and have immediate public access, at low cost. Depositors prepare their files and metadata. When data are published, depositors get a DOI and a data citation. openICPSR curatorial experts conduct a metadata review and add the dataset information into the ICPSR data catalog, where they are searchable by researchers and students at ICPSR's 750 member institutions, as well as by the general public.



The openICPSR homepage is designed so users can begin to deposit their data easily.

• The Professional-Curation Package provides full metadata generation and bibliography search, statistical package conversion, and user support. Because the depositor fee depends on the data complexity, depositors need to contact ICPSR for pricing.

• The Full Topic Archive Package enables an institution — such as an agency, foundation, university, or department — to disseminate data publicly. It provides professional curation as well as premium services such as dedicated staff specialists, a website, custom data tools, data community outreach, and acquisitions and compliance reporting. Project managers, officers, and agencies need to contact ICPSR for a proposal.

All openICPSR deposit options provide bit-level preservation and public access for the long term.

Significantly, openICPSR also provides researchers with the ability to make their restricted-use data publicly available.

"We accept sensitive data and make it public in an application-type environment," said Detterman. "We'll store it and disseminate it securely to the public via our Virtual Data Enclave, and we have the knowledge of how to do that — not everyone does.

"We're charging a fee to sustain openICPSR so the data held in it are available not just today, not just tomorrow, but into the future," she

Wherever possible, the costs of archiving should be built into research grants to avoid burdening the researcher or the depositor's institution.

Linda Detterman,
ICPSR director of Marketing and Membership

Service */ Continued on page 3*

Service */ Continued from page 2*

said.

"Curation carries costs; there are professionals involved," she added. "There's the staff time of our librarians and metadata specialists, who are reviewing the data and cataloging it, and we have the technical professionals, both the website and the individuals who are programming and maintaining the site functionality. There also are costs for multiple copies of the datasets, and server costs, and many others."

Wherever possible, the costs of archiving should be built into research grants to avoid burdening the researcher or the depositor's institution, she added.

Alter: openICPSR helps researchers meet public-access requirements

"Our competitors in Europe have decided that research data are essential scientific infrastructure, and they are funding data repositories directly. NSF wants the data to be open access, but they expect researchers to do it themselves," Alter said. "So Principal Investigators should include funds in their research grants to cover depositing data into openICPSR."

While openICPSR provides a new trusted and sustainable service to address open-access requirements, it is not a replacement for membership in ICPSR, Detterman said. Member deposits to ICPSR are thoroughly curated to correct for missing or misleading documentation, missing values, corrupted files, and other problems. Moreover, members have access to over 28,800 members-only datasets, discounted tuition for the ICPSR Summer Program courses, and other exclusive services.

An ICPSR task force with members from the Data Acquisitions, Computing and Network Services, Curation, Web Team, and Marketing and Membership units planned and guided development of the service. Prototyping, programming, and testing were handled by skilled professionals in-house.

"The data-sharing environment is changing all the time, and openICPSR provides a worthy complement to ICPSR membership, in providing researchers, institutions, and funders the data sharing solutions that serve their various purposes," Detterman said.

Mark Thompson-Kolar can be reached at mdmtk@umich.edu.

Learn more about openICPSR

- Explore www.openICPSR.org
- Find our presentations on www.slideshare.net (search: "openicpsr")
- Download the PDF brochure from our website
- View our openICPSR instructional videos at www.youtube.com/user/ICPSRWeb:
  - Welcome to openICPSR
  - Orientation to ICPSR's Public Access Data Collection
  - Preparing to Upload Files to openICPSR
  - Uploading and Describing Files on openICPSR

Submissions sought for Data Curation Student Research Paper Competition

ICPSR invites original student submissions for a research paper competition on data curation.

The competition highlights exemplary student research on data curation, including but not limited to topics such as data management planning, support of the data curation lifecycle, metadata, data confidentiality, preservation, and cost modeling.

It is open to all undergraduate, master's, or PhD students, and recent graduates who graduated on or after April 1, 2013. Entrants may be from the US or outside the US.

The first-place winner receives $1,000; second place, $750.

The submission deadline is May 30, 2014. For details and entry forms, visit the ICPSR paper competition web page.

Summer Program course topics include curation practices, teacher effectiveness data, transparency, analysis workflows

More than 80 courses are offered, including a variety of new workshops

The 2014 ICPSR Summer Program in Quantitative Methods of Social Research will offer more than 80 courses in statistical and methodological topics ranging from beginner to advanced.

As in prior years, some training will be provided in locations other than Ann Arbor (see inset).

This summer's offerings include several new workshops on a variety of analytical and substantive topics, as well as returning workshops in social science disciplines that are relatively new to the Summer Program curriculum. These workshops include those listed below, which will be held in Ann Arbor unless otherwise indicated. (Please note the application deadline has passed for some workshops.)

Transparency Practices for Empirical Social Science Research, June 2-6, will explore tools and techniques that increase transparency in research, including systematic disclosure of methods and results, registration and pre-analysis plans, and open data and materials. Location: Berkeley, CA.

The Measures of Effective Teaching Longitudinal Database: A Review of the MET Project and Available Data, June 9-11, will discuss key elements of this project and resulting database, including nested data structure (district, school, teacher, student), randomization process and implications for analysis, and student surveys. No tuition for accepted participants.

Introduction to Survey Methodology: Questionnaire Design, Data Collection Modes, & Improving Survey Quality, June 9-13, will examine research literature as well as experiences in designing, conducting, and analyzing surveys, including inference and error, modes of data collection, interviewing, questionnaire design, question evaluation, item construction, and nonresponse.

Analyzing Intensive Longitudinal Data: A Guide to Diary, Experience Sampling, and Ecological Momentary Assessment Methods, June 24-27, will explore methods that allow the study of people's thoughts, emotions, and behaviors in their natural contexts, usually via self-reports.

About the 2014 ICPSR Summer Program

The Summer Program will offer 85 courses taught by 111 instructors.

Four-week sessions: 42 courses

- 28 statistical workshops
- 11 lectures
- 3 Blalock lectures

Short workshops: 43 courses

- 20 statistical workshops in Ann Arbor
- 5 substantive workshops in Ann Arbor
- 18 statistical and substantive workshops in other locations (Amherst, MA; Berkeley, CA; Boulder, CO; Chapel Hill, NC; Manhattan, NY; and Montreal, QC, Canada)

Location: Amherst, MA.

Bayesian Methods for Prevention & Intervention Science, June 30-July 2, will introduce basic elements of Bayesian statistics and explore relative advantages of the Bayesian perspective over the frequentist perspective. No tuition for accepted participants.

Item Response Theory: Methods for the Analysis of Discrete Survey Response Data, June 30-July 3, will introduce methods useful in the development and analysis of survey measures (political, sociological, educational, test scores, etc.) in which item responses are in the form of discrete categories.

Managing Statistical Research: The Workflow of Data Analysis, July 7-11, will explore how to plan, organize, document, and execute sophisticated quantitative analyses (regardless of the statistical methods used) so that you work efficiently and accurately while producing results that are replicable.

Finding Patterns in Data, Big and Small, July 21-25, will focus on finding groups in "regular" datasets, network data, longitudinal data or repeated measure data, text mining, or proximity data using cluster analysis, algorithms, and formal statistical models.

Empirical Modeling for Theory Evaluation, July 21-August 15, will explore how to specify, estimate, and interpret empirical models of the complex context-conditionality and the ubiquitous temporal dependence (as well as spatial interdependence

Summer */ Continued on page 5*

***Summer** / Continued from page 4*

and endogeneity), which characterizes modern, sophisticated social science theory.

Designing, Conducting, and Analyzing Multi-Racial and Ethnic Political Surveys, July 28-30, will focus on developing multi-racial/ethnic political surveys including collaborative funding strategies, modes of data collection, and appropriate research designs.

The Pathways to Desistance Study: Analyzing the Life Event Calendar Data for Substance Abuse Research, July 28-30, will explore methods and measures of this longitudinal study of serious adolescent offenders as they transition to adulthood and out of crime. No tuition for accepted participants.

Curating & Managing Research Data for Re-Use, July 28-August 1, will introduce efficient curation practices to ensure the usability, quality, and safekeeping of data resources.

Providing Social Science Data Services: Strategies for Design and Operation, August 4-8, will touch on the complete research lifecycle. Fundamental topics are covered, including conducting a data reference interview, searching for social science data, interpreting documentation, coping with various dissemination formats, accessing and subsetting data through web-based tools (e.g., SDA and Nesstar), selecting and downloading ICPSR data, and options for local data delivery.

Latent Class Analysis in Social Science Research, August 11-15, focuses on this statistical method for identifying unobservable or unmeasured subgroups within a population.

For more information, contact the Summer Program at 734-763-7400 or sumprog@icpsr.umich.edu.

Jacoby steps down from Summer Program directorship

ICPSR Director George Alter, left, praises William Jacoby's accomplishments as director of the ICPSR Summer Program during a farewell event in December. Jacoby, center, left the position to become editor of the American Journal of Political Science. John Garcia is serving as interim director as a search for a new permanent director goes forward. At right is Dieter Burrell, Summer Program assistant to the director for academic support services. (Photo by Mark Thompson-Kolar)

3 new webinars available on resources at ICPSR

ICPSR conducted the following webinars in recent months to inform users about ICPSR resources.

- **"Resources for Health Research at ICPSR,"** presented by directors of specialized archives with health research data: slides (PPTX, 7.9MB), video.
- **"Guidelines and Resources for OSTP Data Access Plans,"** presented by Jared Lyle and Amy Pienta: slides (PPTX, 5.5MB), video.
- **"2014 Student Research Opportunities at ICPSR,"** presented by Abay Israel, who explained ICPSR's Summer Internship program and Research Paper Competitions: slides (PPTX, 17 MB), video.

Number of projects using Measures of Effective Teaching data grows

Over 150 researchers on 34 teams now utilize MET Longitudinal Database's classroom videos, datasets

By Mark Thompson-Kolar
ICPSR Editor

Just over seven months after ICPSR began accepting applications from the research community at large for use of the Measures of Effective Teaching Longitudinal Database, 155 researchers are working with the data on dozens of projects.

The researchers are members of 34 teams making use of the MET LDB quantitative and/or qualitative data, said Johanna Bleckman, an ICPSR manager of the database project. The MET LDB provides archival, dissemination, and training services for the Bill & Melinda Gates Foundation-supported MET Project, the largest study of classroom teaching ever conducted in the United States. Additionally, about 25 teams are in the process of applying for access to the data.

"Initially, the data were available only to the MET Partners — the researchers and organizations who handled various aspects of the original data collection — and to a set of 10 winners of the MET Early Career Research Grants, which were announced last year," said Bleckman. "But now that the MET LDB is available to the entire research community, we're seeing the interest spread. The studies that are using the MET LDB data are pretty cutting edge."

The MET LDB holds about 60 datasets containing a variety of indicators of teaching quality collected in classrooms of more than 2,500 4th- through 9th-grade teachers in six US school districts and over 300 schools. Information about these restricted-use data is searchable on a dedicated website and accessible to approved researchers in a Virtual Data Enclave at ICPSR that provides secure online access. Documentation files are available for public download.

measures of effective teaching
LONGITUDINAL DATABASE

Read stories of three projects using MET Project data. **Page 7**

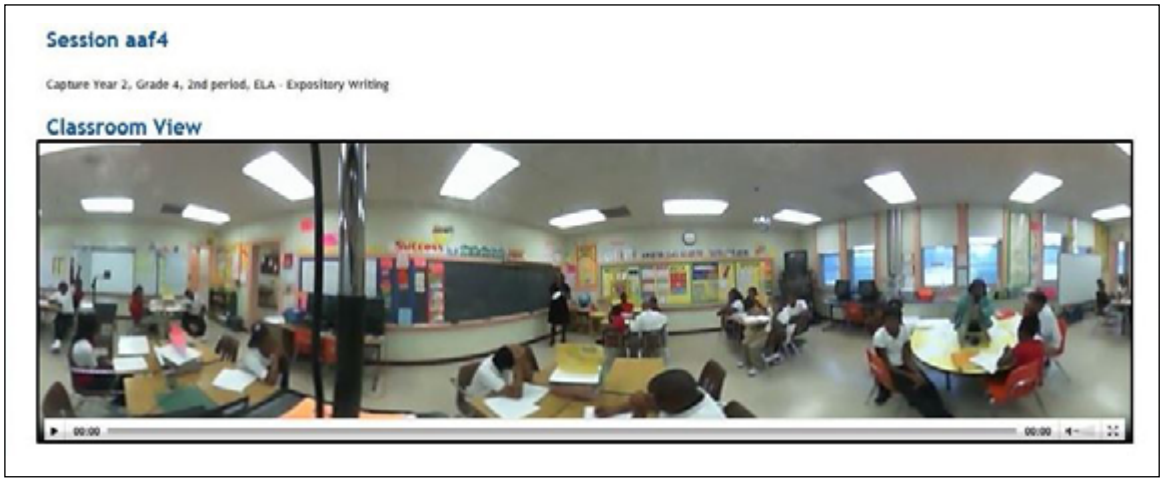
Bleckman: Studies using the data are on cutting edge

Jekielek: The data are qualitative and quantitative

Importantly, the MET LDB provides access to an abundance of qualitative data in the form of more than 11,500 videos recorded in 1,424 teachers' classrooms. These restricted-use videos are accessible via a secure streaming player developed by ICPSR for use with the MET Project videos.

"Project investigators were interested in understanding how classroom observation along with teacher and student assessments can be used to evaluate teaching, and then creating reliable and valid teacher evaluation systems, in addition to training programs, based on those findings," said

Teachers */ Continued on page 7*



A screen shot from the MET LDB shows a scene from a classroom video. The database contains more than 11,500 videos. (Courtesy MET LDB)

***Teachers** / Continued from page 6*

Susan Jekielek, Education and Child Care Data Archive director. "The fact that MET is the largest study of classroom instruction, and that the data are a combination of quantitative files along with thousands of hours of videos really makes this a standout study for researchers to work with."

A few weeks ago, the MET LDB holdings were expanded to include two additional years of districtwide student standardized test scores from all six of the school districts, bringing the total number of years to four. These data, obtained through agreements established by ICPSR with the districts, can be used to create value-added measures to attempt to quantify the value a teacher added to student's knowledge or education over the course of a year.

"Researchers can look longitudinally at students in order to estimate the effects of a particular teacher," Bleckman said. "Access to this kind of districtwide data is very valuable and difficult to obtain."

Sources of information about the MET Project

- Visit the MET Project website
- Visit the MET LDB website
- Email us at met-ldb-inquiries@umich.edu

The possibility exists that more districtwide data could be introduced into the database in the future, further enhancing its research value, she added.

Other data also were added to the MET LDB over the fall and winter, as a result of requests from researchers starting in late 2012. "We received feedback from users that they need data we didn't think were going to be important to include in the database," she said. "We've been able to process and release them in addition to the original suite we anticipated releasing." Several files — including data that can help researchers interpret primary research data — were approved for inclusion by ICPSR directors and the Gates Foundation.

An example, Bleckman said, is the set of files containing master-coded scores for video ratings, allowing researchers to evaluate the consistency and efficacy of the raters and rating systems. Another example is a pair of Excel files containing comments from video raters.

The MET Project and data in the LDB are providing vital insights into what constitutes high-quality teaching, Bleckman said. "This project and the sharing of its data make it possible for researchers to further quantify good teaching — to understand it as a science that can be identified, evaluated, and systematically improved."

Mark Thompson-Kolar can be reached at mdmtk@umich.edu.

measures *of* effective teaching
LONGITUDINAL DATABASE

Researchers obtaining important benefits from MET Project data

By Mark Thompson-Kolar
ICPSR Editor

Researchers already using the Measures of Effective Teaching project data are finding them an excellent resource. In 2013, the University of Michigan's Institute for Social Research and the National Academy of Education announced that 10 educational researchers had received MET Early Career Research Grants, providing them with a year's access to the MET Longitudinal Database (MET LDB) and $25,000.

Here is an overview of three grant recipients' projects.

Tanner LeBaron Wallace

Tanner LeBaron Wallace, an assistant professor of applied developmental psychology at University of Pittsburgh, is studying adolescents' reaction to the MET Project classroom videos. She shows adolescents sections of classroom videos and then scores their reactions in order to parse different influences on the classroom environment.

"I suspected that if you paid attention to what adolescents were saying about classrooms,

***MET** / Continued on page 8*

Wallace: Research includes reactions of adolescents

***MET** / Continued from page 7*

you might be able to figure out what they are paying attention to that's not currently picked up in the adult-developed observational protocols around managing classrooms," Wallace said. Adolescent participants view the same MET videos that already have been viewed and coded by adults, and their verbalized observations are recorded and provided time stamps for comparison with the adult video coders' notes.

"A challenge has been that at first, the adolescents were commenting on the most simplistic things, like, 'That student's not paying attention.' In my other data-collection efforts, I knew that adolescents had really rich theories about effective teaching and interaction," she said. To ensure the research generated meaningful data, her team developed a protocol to garner more useful responses to the videos, the Adolescent Affective Response Rating, which will be utilized for a large number of adolescent participants in late spring.

Having adolescents viewing the MET videos presented potential confidentiality issues for Wallace's project, and ICPSR was able to help solve them. "The adolescents are actually viewing MET data, so we needed a confidentiality agreement just like any research team member would sign," Wallace said. "But they're not a member of my research team; they're a study participant. Johanna (Bleckman, an ICPSR manager of the MET LDB project) really helped everyone understand what I had been funded to do and what it required, and how that would actually work."

ICPSR had the resources to provide a Participant Pledge of Confidentiality for Wallace's project, Bleckman said. With it, adolescent participants in a research project "agree to abide by the same rules that the Principal Investigator and all of the staff are, but the participants are being shown the videos in a totally different capacity." ICPSR's extensive experience with restricted-use data enables it to provide a comprehensive, cutting-edge dissemination process — legally, ethically, and technologically.

Wallace said Bleckman also provided invaluable assistance by helping navigate another video consent-related concern during the IRB approval process. Because it's restricted-use data, "we had to work to identify a video to show (to the IRB) that had been consented in a way that it could be shared with someone who wasn't part of my research team."

The large number of files provided by the MET LDB brings a bounty of opportunities, Wallace said. "To collect video is so intensive and requires so much work — so the fact that I can have access to thousands of videos that I did not have to collect is such an opportunity for me. I had never thought the kind of research that I do could ever have the ability to develop protocols that can then generate variables that I can actually statistically model. It's really exciting."

The research has the potential to round out theories of effective teaching, she said, by incorporating adolescent interpretations of classroom interactions to improve existing observational protocols. "It has huge implication for teachers' professional development."

Rachael Gabriel

Rachael E. Gabriel, an assistant professor of reading education at the University of Connecticut, and her three Co-PIs are investigating the relationship between validated measures for identifying teacher effectiveness and existing research on effective literacy instruction in English Language Arts.

Her team is using the MET LDB as a source of videos that are already attached to validated ratings of teacher effectiveness. "We're using the video database because we're rescoring the videos and transcribing them to be analyzed. We're also using from the database itself teacher value-added data and teachers' ratings on the three evaluation protocols that the MET study used," she said.

Gabriel: Team is using MET LDB as source of videos

Value-added assessment measures teachers' educational contribution in a year by comparing their students' test scores with those of the same students in prior years, and to scores of other students in the same grade.

"We're trying to see what kind of information about teaching we gain or lose when we 'zoom way in' and do a much more detailed job than any of the protocols MET used, and also 'zoom way out' and do a much higher-level checklist approach," she said. "We're also seeing if taking those two contrasting approaches can let us connect the same way with value-added measures as the (other protocols) have."

Gabriel said the size of the MET LDB offers major advantages. "One of them is just how big it is: there's a lot of data in it, and another is that it already has that value-added data, which is used as a gold standard in validating observation tools," she said.

"We want to make sure that we

***Projects** / Continued on page 9*

Projects */ Continued from page 8*

are answering the questions that will come up about value-added data, and that kind of data is time-consuming and expensive to generate," she said. "The fact that it exists already and we have videos that we can look at through our own lenses but we still know what value-added has to say about them is really unique." The size of the database also allows it to represent many grades and provide multiple observations per year, she added.

"The MET database is really important because it lets us hold on to a representation of practice that isn't bound to some sort of rating system; we still have the video to go back to," Gabriel said. "It isn't just a database of scores that we generated in different ways; we have the video record of what went on. So as our understandings of what best practices are change, we'll still have this connection between the test scores, the value-added data, and what we see."

The other Co-PIs on the project are Morgaen Donaldson, Sarah Woulfin, and Kim LeChasseur, all of University of Connecticut.

Matthew Steinberg

Matthew P. Steinberg, an assistant professor of education at the University of Pennsylvania, and Co-PI Rachel Garrett, of the American Institutes for Research, are exploring whether the effect of observed instructional practice varies by student characteristics and how teacher instructional practice might impact students who differ on observable characteristics such as poverty, prior achievement, and status as an English language learner.

"For policy purposes, we wanted to know to what extent — based solely on a teacher's observation score from their classroom lessons — can we say something causal about teacher effectiveness," Steinberg said. "With the MET study, we're presented a really nice opportunity to say something about causally identifying effective teachers."

Among the team's early findings is that from year to year, the correlation of a teacher's observation score is limited, at about 0.5, "which means much of what we're observing about teacher practice is very context specific," he said. "For those of us who have been in classes, we know that the idiosyncratic makeup of a classroom very much shapes the type of instruction that takes place."

Steinberg: Team examining results of student sorting

The team's first paper, "Examining Teacher Effectiveness using Classroom Observation Scores: Evidence from the Randomization of Teachers to Students," is under final review at a journal, Steinberg said.

The team now is examining pathways of student sorting — what types of results occur when students do not comply with random assignment to a teacher and instead end up with a different teacher. "In a natural setting, we observe the classes that kids end up in and the teachers that end up in those classes, but what the randomization the MET study allows us to do is to say, 'Using the randomization as sort of a baseline, let's now look to see how students and teachers are resorted, and let's look to see the kids that these resorted kids end up with: Do they end up with higher-achieving or lower-achieving kids? Let's look at the characteristics of the teachers that the kids end up with when they're moved away from — or when they're kept with — the randomized teacher,' " he said.

"It seems to be that there is some adverse effect of resorting kids. That calls into question in many ways what we all believe, which is that there is this kind of positive sorting that happens in schools, where principals or teachers are matching the highest-performing kids with the highest-performing teachers, and that has a positive effect on those kids' performance. These are very preliminary findings; we're still trying to sort this out."

The project utilizes the MET teacher data, including teacher observation scores and value-added scores, and then links those teachers to their students in the second year of the MET study, 2010-2011. "We've got a rich set of student characteristics; we have student test scores, and what's principally important is being able to say who the randomized teacher was within grade-subject combinations," he said. The team is not using the MET LDB videos.

Steinberg praised the MET project's approach. "What was almost a coup from a researcher perspective was getting the randomization in the second year of the MET study, and allowing researchers like myself to really try to better understand whether we can causally identify effective teachers in a world where much of the research we do is observational, it's non-experimental."

Mark Thompson-Kolar can be reached at mdmtk@umich.edu.

ICPSR Council and staff members address ICPSR business during a meeting on March 6 in the Perry Building.

New ICPSR Council members have first meeting together

By Mark Thompson-Kolar
ICPSR Editor

The ICPSR Council met March 6-7 in Ann Arbor, with all six newly elected members in attendance.

New Council Chair Christopher H. Achen of Princeton University led the meeting.

The new members are:

- Janet Box-Steffensmeier, Ohio State University
- Robert Chen, Columbia University and Center for International Earth Science Information Network (CIESIN)
- Philip N. Jefferson, Swarthmore College
- Chandra Muller, University of Texas at Austin
- Ronald Nakao, Stanford University Libraries
- William Vega, University of Southern California

Biographical sketches and photographs are available on the Council web page.

Past Chair Rogelio Saenz of the University of Texas at San Antonio will continue on the Council for one year.

The new members received an orientation and met in their assigned committees for the first time.

Agenda items at the March meeting included the budget; implementation of the ICPSR Strategic Plan; the Cash Reserves Policy; approval of ICPSR Bylaws; the Policy on Codistribution of ICPSR member-funded data; an update on openICPSR; the future of the Replication Data Archive; information about ICPSR's Government Data Strategy; a report on the transition of the Summer Program director and interim director; an update on Summer Program scholarship and fee-waiver initiatives; a Summer Program report on e-courses and MOOCs; and updates from Council committees.

At the end of the meeting, the Council approved action items including an update to the Cash Reserves Policy, revisions to the Bylaws, an amendment to the Codistribution Policy, and a request for the staff to develop criteria for Summer Program scholarships for under-resourced institutions.

New Council members William Vega, left, and Philip N. Jefferson listen during a committee meeting.

(Photos by Mark Thompson-Kolar)

Data from '50s Cornell Study of Occupational Retirement available

Historic study was migrated from punched cards to modern digital format as part of an ICPSR project

By Kate Billerbeck
For ICPSR

ICPSR recently released the data from the Cornell Study of Occupational Retirement, 1952-58, through the National Archive of Computerized Data on Aging (NACDA).

Originally conducted in the 1950s by Principal Investigators Gordon F. Streib, Wayne E. Thompson, Milton L. Barron, and Edward A. Suchman, all of Cornell University, the study aimed to collect and analyze information about the then-"well-defined" life transition from employment to retirement in America.

The study followed for six years a cohort of 4,032 workers who were 64 years old in 1952. The workers were employed at over 250 companies and government agencies across the United States.

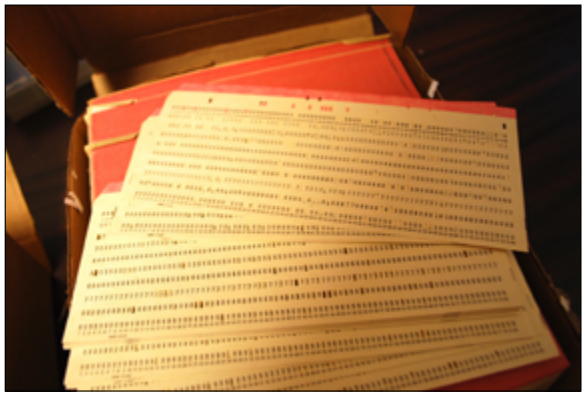
Following the study's completion, its materials were stored in seven filing cabinets at the Newberry Library in Chicago and were given to ICPSR as a gift in the mid-1990s.

"Gordon Streib was one of those really important social gerontologists, one of the giants of the field," said Amy Pienta, ICPSR's director of Data Acquisitions.

Restoration of these data for public was an ongoing process for ICPSR's Acquisitions group. Because the data had not been analyzed for over 30 years, none of the information had been digitized before ICPSR received the study materials. Original data

Streib */ Continued on page 12*

This punched card reader was purchased and repaired by ICPSR for the project. Below, the hardcopy materials were stored for decades in file cabinets. (ICPSR Photos)



An assortment of punched cards is shown at left. The process of converting the data from physical media to digital files required years of careful work.

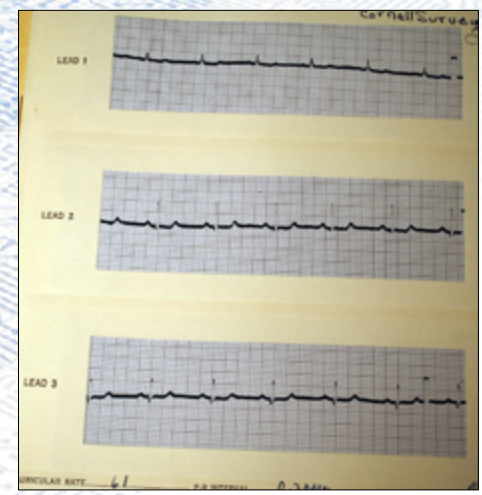
***Streib** / Continued from page 11*

were stored on punched cards, and questionnaires and supplementary materials, such as participant medical records and reports, existed only in hard-copy form.

Pienta is the PI on the data-recovery project. ICPSR's director of Curation Services, Jared Lyle, also contributed to the project, along with research assistants Kate Billerbeck and Joanna Tatomir.

Streib visited ICPSR in 2003 to help the project members understand the data, Pienta said. "He gave us the history of the study and why it was important, from a scientific point of view. This was 40 years after he had collected it. It was great to have him here for that." He died in 2011 at age 92.

In order to restore the usability of the data, the punched cards were read by a Cardamation punched card reader (purchased by ICPSR) and translated to a contemporary statistical format. This process used punched-card reader code supplied by the Roper Center for Public Opinion Research, a partner on one of our projects, Pienta said.



The Cornell Study of Occupational Retirement data include health information, as shown on this electrocardiogram record. (ICPSR photo)

The newly migrated data were cleaned and processed. The study's codebooks also were digitized.

The files include baseline data and data from four follow-up waves (ICPSR 34918). All are available for public use.

Through its longitudinal design, the Cornell Study of Occupational Retirement provides a unique perspective on retirement in the 1950s. The data cover gender and health information — aspects of retirement that were seldom considered during that time period — and their relationship with this life transition.

"These are phenomenal data," Pienta said. "The time period of the '50s is important because Social Security legislation had recently been implemented – so these are some of the earliest cohorts to make retirement decisions with the new rules shaping their careers. That's historically significant."

Moreover, the original analyses using these data would not have included multivariate modeling techniques because they didn't get applied to punched card data, she said. "Having this study migrated to modern digital formats presents a nice opportunity to take this really large-scale set of data and apply some of those techniques to look at how retirement in the 1950s and '60s differs from what happened in the 1970s and onward."

The recovery of the data was funded by the National Institute on Aging (1 R03 AG04038).

ICPSR Senior Editor Mark Thompson-Kolar contributed to this article.

Updated Strategic Plan website launched; projects related to directions under way

ICPSR's new Strategic Plan is now available online. The Plan is the result of about a year of effort to develop a multiyear vision and set priorities for the organization to maintain its leadership position in data curation and stewardship and to advance social and behavioral research.

The Plan lays out three strategic goals and four strategic directions

for ICPSR. These directions and their related strategies leverage the organization's historic and current success, its strong membership and partnership network, and its position as a leader in order to increase stakeholder value and support a vibrant field of social and behavioral research. The Plan also introduces an updated Mission Statement.

Since the Plan's launch, the ICPSR staff has begun working on initiatives corresponding to the strategic directions.

NEWS AND NOTES ...

CDC Scholars Program grows

The Visiting Scholars Program of the China Data Center (CDC) at the University of Michigan is hosting a record 16 participants in the current academic year. The increase comes on the heels of the highly successful class the prior year, which had nine visiting scholars, up from three in 2011-12.

Module teaches data with love

ICPSR's Resources for Instructors recently released a new Data-Driven Learning Guide, "Is Love Really Blind?" DDLGs are designed by teaching faculty to help build data literacy in undergraduate classrooms. This DDLG helps students explore factors that influence the development of romantic relationships using crosstabulations and comparison of means. ICPSR now offers 52 DDLGs covering a broad range of social science topics.

ICPSR moves to DataCite for DOIs

ICPSR changed its DOI registration agency to DataCite at the end of 2013. Existing DOIs were unaffected. ICPSR is an associate member of DataCite.

More restricted-use data details

In an effort to provide more detail about handling of restricted-use data, the ICPSR website's Restricted-use Data Management area was expanded with additional content. It now contains explanatory sections about identifying, accessing, and depositing restricted-use data; as well as about ICPSR's disclosure risk review.

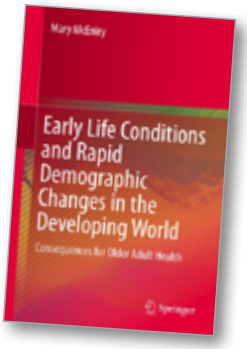
China Data Center receives grant

ICPSR partner the China Data Center at the University of Michigan and the Center on Religion and Chinese Society at Purdue University announced the "Spatial Study of Chinese Religions and Society" project, supported by a three-year, $400,000 grant from the Henry Luce Foundation. The project promotes empirical, qualitative, and quantitiative research on Chinese religions.

White paper urges new approaches to assure access to scientific data

More than two dozen data repositories serving the social, natural, and physical sciences this winter released a white paper recommending new approaches to funding the sharing and preservation of scientific data. "Sustaining Domain Repositories for Digital Data: A White Paper" (PDF 52KB), is an outcome of the meeting convened June 24-25, 2013, in Ann Arbor. The meeting was organized by ICPSR, supported by the Alfred P. Sloan Foundation, and attended by representatives of data repositories from a range of scientific disciplines.

Book explores early life conditions



Researcher Mary McEniry, director of the Data Sharing for Demographic Research project at ICPSR, has published "Early Life Conditions and Rapid Demographic Changes in the Developing World: Consequences for Older Adult Health." The book examines how rapid reductions in child and infant mortality from the 1930s to the 1960s without an accompanying rise in standard of living in some countries resulted in cohorts that included larger numbers of infants and children who survived exposure to poor nutrition and infectious diseases but whose health may now be more at risk as they age because of these early life conditions.

Article provides research insights

The Resource Center for Minority Data at ICPSR has made available a new article based on an in-depth interview with Ruben G. Rumbaut, professor of sociology at University of California, Irvine, and a Principal Investigator of the Children of Immigrants Longitudinal Study (CILS). Rumbaut explains how the CILS began and comments on its objectives and major findings — providing useful advice for other researchers who are analyzing the data.

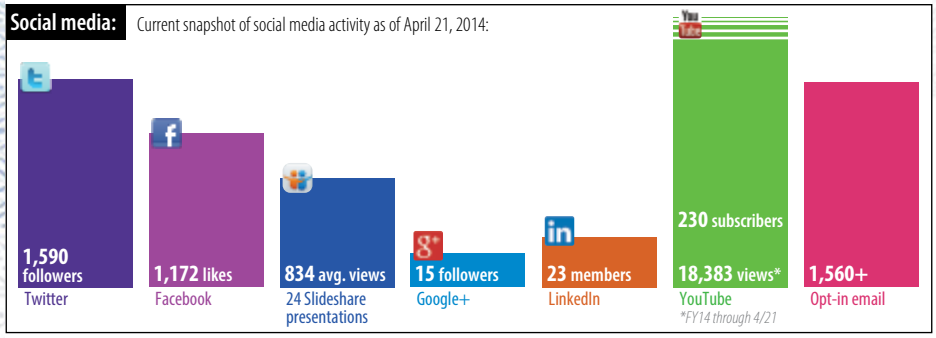
Proposals sought for NCAA studies

The NCAA is seeking proposals for studies of issues important to NCAA student-athletes and NCAA member institutions using data from its Growth, Opportunities, Aspirations, and Learning of Students in College Study (GOALS). These data are housed in the NCAA Student-Athlete Experiences Data Archive at ICPSR. In 2006, GOALS survey responses were received from more than 19,780 student-athletes at 620 NCAA member institutions. Respondents provided a range of information about their lives as student-athletes. In the first phase of this pilot program, the NCAA will approve data access for a limited number of proposals.

NEWS AND NOTES ...

ICPSR Grows Online

Social media: Current snapshot of social media activity as of April 21, 2014:

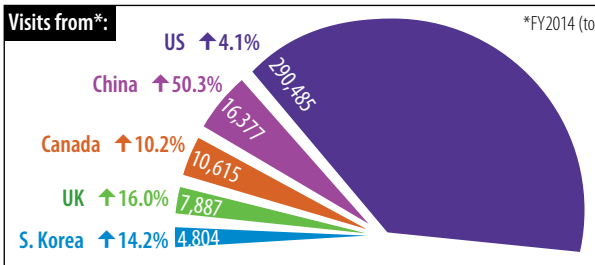


Website activity: FY2014 (to April 21, 2014) vs. same period in FY2013:

248,560
# of people who have visited icpsr.umich.edu

Sessions (visits)	378,430	↑6.4%
Users (unique)	248,560	↑5.1%
Pageviews	4,283,308	↓1.9%
Pages/Session (visit)	11.32	↓7.8%
Avg. Session (visit) Duration	9:52	↑2.8%
% New Sessions (first-time visits)	57.8%	↓1.5pp

Visits from*: *FY2014 (to April 21, 2014). Changes are vs. same period in FY2013:



Top Referrers:
NCES
UNC
zhidao.baidu.com (China)
MIDUS – Wisconsin-Madison
Princeton

Upcoming opportunity!

Save the Date — 2014 Data Fair

October 7-9, 2014

The 2014 ICPSR Data Fair, a series of webcasts featuring sustainable data access, will take place October 7-9.

The fair will cover topics including an orientation to federal data sharing requirements, recommendations for how to achieve good data stewardship, valuable resources for creating data management plans, and tips for evaluating data sharing services. The fair is also expected to feature a number of public-access data collections.

The Data Fair is designed for the social and behavioral sciences community at large, including researchers, librarians, teaching faculty, students, and policy makers. The event is open to everyone and will use readily available, free technology that allows access to the webcasts without any software-download requirements.

Please save the dates on your calendar, and check the ICPSR website for updates on details of the 2014 ICPSR Data Fair!

INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH

George Alter, Director
Peter Granda, Associate Director
Mary Vardigan, Assistant Director
Diane Winter, Assistant Director for Administration

Council Members

Christopher H. Achen, Chair
Princeton University

Marilyn Andrews
University of Regina

Janet M. Box-Steffensmeier
Ohio State University

Tony N. Brown
Vanderbilt University

Robert S. Chen,
Columbia University & CIESIN

John Fox
McMaster University

Philip N. Jefferson
Swarthmore College

Carl Lagoze
University of Michigan

Chandra L. Muller
University of Texas at Austin

Ronald Nakao
Stanford University Libraries

Rogelio Saenz, Past Chair
University of Texas at San Antonio

William Vega
University of Southern California

Linda J. Waite
University of Chicago

Bulletin Staff

Editor: Mark Thompson-Kolar
mdmtk@umich.edu, (734) 615-7904

ICPSR is a center within the Institute for Social Research at the University of Michigan.

INSTITUTE FOR SOCIAL RESEARCH
UNIVERSITY OF MICHIGAN