Jessica M. Sautter, *University of the Sciences*

CUR Focus

Secondary Analysis of Existing Data in Social Science Capstone Research

Secondary analysis of existing data is an important model for capstone research in social science because it allows students to integrate and apply the knowledge gained through coursework in their majors or minors. Students identify connections among concepts learned through coursework, identify appropriate data and statistical methods to answer relevant research questions, and use statistical modeling to demonstrate and interpret relationships among concepts. Through activities that include a research proposal, a proposal to the campus institutional review board (IRB), an annotated bibliography, data analyses, and a final public presentation, students gain applied analytical skills and make progress toward mastering the undergraduate learning goals of self-directed learning, critical thinking, scientific reasoning, ability to conduct statistical analysis, and effective oral and written communication.

The advantages of using the secondary-analysis model in capstone classes include efficiency, publishable outcomes, transferrable skills, and student-faculty collaboration. This model is especially effective for students pursuing careers in the health sciences because it provides a complementary approach to understanding health problems and develops skills and content knowledge necessary for success in medical or professional school or in analytical careers.

Specifically, secondary analysis of existing data involves obtaining datasets from studies that have already been completed and using the data to answer new questions (Smith et al. 2011; Trzesniewski et al. 2010). The major advantage of this research model is that it eliminates the time and expense of gathering data and relies on high-quality, reliable data collected by experts. The major disadvantage is that the undergraduate researcher has no control over the contents of the data, so she or he must search for datasets appropriate to answer particular research questions. Secondary data can be physical (e.g., lab specimens), qualitative (e.g., in-depth interview transcripts), or archival (e.g., newspaper contents). Most secondary datasets are based on large-scale surveys, are intended to be nationally representative, and utilize quantitative self-reported data. These data can cover the social, demographic, economic, and other characteristics of individuals; behaviors, attitudes, preferences, and beliefs; and, sometimes, administrative, physical, and genetic data. Two examples of general social science datasets that are often used for secondary analysis are the U.S. General Social Survey and the global World Values Survey.

The secondary-analysis model fits all of the essential features of good undergraduate research, educating students in reading scientific literature, developing research questions, designing studies, working independently and collaboratively with faculty, mastering research techniques, improving oral and written communication, and providing opportunities for course credit, presentation, and publication (Lopatto 2003). This model of undergraduate research is based in student-centered, active-learning methods (Blumberg 2005). As a form of inquiry-based learning, it can increase students' science-literacy skills and self-efficacy (Brickman et al. 2009), which are essential for the growing number of employment opportunities that require quantitative analysis. Although the outcomes are clear, the process can be difficult for both students and faculty. The model, methods, advantages, challenges, and applications presented in this article will provide a template for students and faculty to use to successfully navigate the intricacies of secondary analysis and maximize learning outcomes.

Model and Procedures

One option for the capstone required for a social science minor at the University of the Sciences in Philadelphia is a course titled Directed Research in Social Science. This particular course requires students to have completed an introductory social science course and a biostatistics course and is recommended as a capstone after they have completed a series of other social science courses. This three-credit course operates on a 15-week semester. Table 1 presents an overview of assignments, learning objectives, and assessments described in more detail below.

Research proposal. Students who elect to pursue directed research typically have a broad idea of topics that they are interested in studying. The first four weeks of the semester involve intensive work identifying specific research questions related to those broad topics. Although this capstone's model focuses on secondary analysis of existing data, this step in the research process also includes analysis of other research methods (e.g., interviews, focus groups, clinical trials), what methods are best suited to answer certain types of research questions, and the advantages and disadvantages of various approaches to the topic of interest. This exercise highlights how different methodologies offer different perspectives on particular topics.

The first step in this process utilizes a concept map to identify broad areas of interest (often the social science

Table 1. Overview of Procedures in the Directed Research in Social Science Course.

Assignment/ Procedure	Learning objectives	Basis of Assessment
Research proposal	Develop informed and feasible research questions and hypotheses based on concept map, literature review, and available data	• Justification of research topic • Clarity and relevance of research questions and hypotheses • Consideration of research designs and available data
Institutional Review Board (IRB) proposal	Design a practical and ethical research project	• Justification of research topic • Accurate description of data source, risk to participants, and protection of personal information • Attention to required details
Annotated bibliography	Summarize, assess, and reflect on scientific literature relevant to the chosen topic	• Thorough literature search technique • Proper citation formatting • Critical review of the literature
Data analysis	Acquire and analyze appropriate data	• Demonstrated understanding of the dataset structure and contents • Accurate and appropriate statistical analysis
Presentation of results	Interpret results and present conclusions	• Organized structure and presentation • Persuasive background and rationale • Clear presentation of data source, measures, and analytic methods • Graphical and textual presentation of results with clear and accurate interpretation • Informative and persuasive discussion

courses that the student has taken), to identify sub-topics within those broad areas, and to make connections among sub-topics. The end result is from one to three research questions that are narrow enough to guide the development of one journal article but broad enough to allow for the revision and adjustment in focus that may become necessary in later steps.

In the second step, students and faculty explore the scientific value of proposed research questions. A brief literature search is conducted to outline what is known about a student's suggested topic, why more research is needed in this area, and what the student's contemplated study can contribute to the knowledge base.

In the third step of this process, a search of available datasets is conducted to assess whether the research questions can be addressed with publicly available datasets. Some common sources of data relevant for social science research include data gathered by the Inter-University Consortium for Political and Social Research, the Pew Research Center, and the federal Centers for Disease Control and Prevention. If no datasets with appropriate variables and research designs are found, then the proposed project is not feasible and students' research questions must be adjusted to reflect the realities of available data.

The students will cycle through the three steps outlined above, with each step improving the feasibility and scientific value of the students' proposed studies. The end result is a three- to five-page research proposal that identifies research questions, justifies the need for the research, and describes the data source(s) appropriate for addressing the research question. Assessment of the students' work is based on their justification for the research topics, the clarity and relevance of research questions, thorough consideration of research designs, evaluation of data sources, and the clarity and organization of the writing in the final research proposal.

Institutional Review Board (IRB) proposal. Protection of human subjects is essential to any research project that analyzes data based on interviews, surveys, or clinical encounters with people. Thus, training in ethical requirements for using human subjects in research and approval of the research proposal by the campus institutional review board (IRB) are important elements of the secondary data-analysis model. Students should complete any institutionally required training in the ethical conduct of research and protection of human subjects (e.g., National Institutes of Health training on Protecting Human Research Participants). In most cases, students will analyze publicly available data that cannot be linked to specific individuals, and they will not directly interact with study participants, so there is no risk of violating privacy protections. Such secondary analysis of publicly available data is likely to be judged as exempt from full IRB review (Puglisi 2001). This is important given that the IRB-approval process can deter faculty from encouraging students' collection of data due to the risk of delays and restrictions (Bledsoe et al. 2007).

The IRB proposal is important for three reasons, however. First, it teaches students about the preparatory work necessary for ethical research. Second, most university IRBs mandate that even research that is likely to be exempt still must be submitted for review, approval, and oversight. Third, formulating the IRB proposal requires students to explain their research goals and methods in a clear and concise manner understandable by outside reviewers. The transition from research proposal to IRB proposal takes approximately two weeks. Assessment of this step is external, based on whether the student explained the project sufficiently well

Secondary Analysis of Existing Data continued

to gain approval from the IRB. Approval is often based on submission of a brief literature review and a justification and rationale for the research project; accurate description of data to be used, risk to participants, and protection of personal information; and attention to required details.

Annotated bibliography. Analysis of secondary data cannot proceed without IRB approval. Therefore, weeks seven through nine can be devoted to further literature review, resulting in an annotated bibliography. Many students reaching the capstone stage were not properly trained or do not remember how to search for scientific literature. Thus, guidance by a librarian well-versed in social science resources would be especially useful at this stage and would lessen the burden on faculty members to mentor students in the search process. This exercise is intended to supplement the original literature review and focus on issues related to selection and analysis of variables. For example, what factors have been demonstrated to be important to the selected research questions? How have previous studies measured those factors? What limitations of previous research should be overcome? What are the limitations of the chosen dataset and measures? Thus, the annotated bibliography includes references that will be cited in the introduction, explanation of data and measurement, and analysis sections of the final research presentation.

Assessment of this procedure focuses on literature-search techniques, citation formatting, and critical review of the literature. The latter should include a summary of each source, assessment of limitations in the reported research and the resulting implications for the proposed research, and reflection on how each source may inform the students' research project (see Purdue Online Writing Lab for more detail: owl.english.purdue.edu/owl/resource/614/01). It is important to emphasize at this point that the literature review phase continues in cycles; after a full draft of the research report is completed, it will be necessary for the student to again search the literature to frame his or her discussion of important findings.

Data analysis. Statistical analysis is at the heart of capstone research modeled on secondary analysis of data. Although a substantial amount of preparation is necessary, this model allows for six weeks of data analysis during a typical 15-week semester. It is crucial that faculty provide adequate guidance on the conceptual, methodological, and applied aspects of statistical analysis, even if the student has recently completed a statistics course. This guidance can take the form of instructional packages covering major areas of analysis, including: (1) creating an analytic sample, (2) inspecting and cleaning data, (3) testing assumptions, (4) descriptive statistics, (5) bivariate analysis, and (6) multivariate analysis. Each instructional package should contain the rationale behind each procedure and a simply-worded guide to conducting analyses and interpreting output for the chosen statistical package. The package also should pose key questions to help the student write about the statistical tests and results.

Several data-analysis packages may be appropriate for student research, depending on institutional resources and faculty and student preferences. SPSS from IBM is one popular package that has a user-friendly point-and-click interface. For students who are more inclined to coding, SAS or STATA is appropriate, especially for weighting with complex sample designs. An open-access option that requires coding knowledge is R. Regardless of the package, students need explicit guides to help them focus on applying their knowledge and answering their research questions. Written guides by the software company and faculty also provide a reference manual to help students apply their analytic skills in later graduate work or employment.

Assessment of the data-analysis component of students' work is based on students' answers to research questions at the end of each written guide. Table 2 demonstrates some typical questions for each step of statistical analysis. Successful completion of this step will lead the student to write most of the measures, analysis, and results sections of a paper or poster.

Presentation of results. Students who complete a research capstone should be required to present their findings, whether through a written term paper, poster presentation, or oral presentation. This final step integrates all of the students' previous work in the semester to produce a final analysis of their social science hypotheses. Assessment is based on how well the student has organized and presented the research report; whether he or she has provided a persuasive statement of the background and rationale for the research; whether the student has clearly presented data sources, measures, and analytic methods; whether graphical and textual presentation of results is clear and the research results are correctly interpreted; and whether the student has provided an informative discussion that persuades the reader that the research was important. The presentation completes the requirements for a capstone course, but students should be encouraged to present their results externally when feasible opportunities arise.

After the Capstone

Students who complete a secondary analysis of existing data should present a poster or oral presentation of their findings at a local, regional, or national scientific meeting, either in social science disciplines or in the student's major

Table 2. A Step-by-Step Guide to Statistical Analysis

Research step	Questions to answer
Creating an analytic sample	• What is the target population of the study sample? • Is the sample representative of the population? • How will you further restrict the sample to match your population of interest? • How many participants have data missing on important variables? • How can you adjust for complex sampling designs?
Inspecting and cleaning data	• What outcome, explanatory, and contextual variables are important? • How are variables coded in the data? • At what level of measurement? • Are the ranges and distributions of variables reasonable and error-free? • What references are needed to determine how to describe and use each variable? • Should you recode variables to facilitate interpretation, fit the study population, or fit established metrics?

Testing assumptions	• What type of statistical models will answer your research questions?
	• Depending on the nature of the outcomes and statistical design:
	  • Are the data normally distributed?
	  • Are observations independent or linked?
	  • Are between-group variances homogeneous?
	  • Are outcome variables modeled by additive and linear combinations of covariates?
	• Do data need to be recoded or restructured to avoid violating assumptions?

Descriptive statistics	• For continuous variables, what is the mean/median and standard deviation/interquartile range? • For categorical variables, what is the percentage and number for each category?
Bivariate analysis	• Are there significant differences in outcome variables by value of explanatory variables? • Do unadjusted regression models show a significant association between each explanatory variable and the outcome variable? • Which variables should be included in multivariate analysis?
Multivariate analysis	• Taken together, which explanatory variables are significantly associated with your outcome variable(s) and how? • How much variation in the outcome does your model explain? • Do the results support or refute your hypotheses?

discipline. This opportunity provides meaningful feedback from new perspectives, adds to the student's experience and resume, and advertises the institution's dedication to training undergraduate researchers. Students should make a serious attempt to publish their work in a peer-reviewed journal, again gaining valuable feedback and experience with the research process while bringing recognition to the student, department, and institution.

Undergraduate research journals are excellent outlets for this work, but ambitious students may wish to submit to a professional journal first. In all forms of presentation, the student(s) should be listed as first author with the faculty advisor as last author. The project is ultimately the student's idea; the faculty advisor's role is to teach students how to properly research that idea and ensure that the final product is of publishable quality.

Learning Goals

Secondary analysis of existing datasets supports the goals for undergraduate learning of critical thinking, scientific reasoning, statistical analysis, and oral and written communication. In one study, while almost all students reported gaining knowledge about a specific topic as a result of their research, social science students particularly valued the learning outcomes of improved communication skills, better problem-solving skills, ability to formulate research questions, ability to contribute new knowledge to society, and opportunity to publish (Craney et al. 2011). Other studies have identified similar learning objectives, including integrating and extending disciplinary content and developing analytical skills and understanding of research (Hauhart and Grahe 2010; McKinney and Busher 2011).

Thus, while learning specific content is important, especially when based on the student's interests, the development of research skills and systematic procedures is also important. Throughout the research process, students must justify their decisions, overcome obstacles, use data to reach conclusions, and effectively communicate their ideas to others. Through this model, students learn a framework to develop methods of inquiry that can be applied to any field or topic in later life. The one-on-one nature of the research capstone provides a safe, guided environment for students to learn valuable job skills.

The secondary-analysis model is rather fixed regarding some definitions of undergraduate research: it is student-initiated, process-centered, and multi- or interdisciplinary; other aspects of the research process, such as level of collaboration

Secondary Analysis of Existing Data continued

and presentation, will vary along a continuum depending on the student's skills and interests (Beckman and Hensel 2009). The model thus allows all students to gain rigorous research experience in a topic that interests them, while allowing flexibility in other areas.

Advantages of the Secondary-analysis Model

One advantage of the secondary-analysis model is that it is *efficient*. Students can identify a topic, conduct a literature review, analyze data, and prepare a final presentation all within one semester. As long as the dataset is publicly available and does not contain information that may identify respondents, the project should be granted exempt status by the campus institutional review board (IRB). Because the data are already collected, students can move right into developing a research question and conducting statistical analyses without the time-consuming and often-thwarted efforts of gaining project approval and recruiting study participants.

Data collection is an important skill worth practicing in many contexts, especially for students with plans for graduate school. However, efficiency in gaining data is a crucial criterion if students are to complete a research experience in the context of a one-semester research capstone course. The secondary-analysis model overcomes major limitations noted in studies of sociology and psychology capstone research—specifically, challenges related to the time needed to collect data and obtain IRB approval (Hauhart and Grahe 2012).

Another advantage of the secondary-analysis model is that it produces *publishable research*. Large, nationally representative datasets, often produced by major universities and government agencies, are viewed as credible sources of data appropriate for publication in well-respected journals. Small datasets designed and collected by students, a common model in undergraduate social science research, are not viewed as having the same quality and rigor and their use may lead to difficulty in publishing the resulting research in peer-reviewed journals. With proper guidance and diligence, it is possible for students and their faculty mentors to submit a manuscript for publication based on as little as one semester's work in a capstone; in most cases, some additional revisions will be necessary after the capstone but before submission.

A third advantage of the secondary-analysis model is that it provides *transferable skills* relevant to many types of research, including clinical medicine, public health, business, and more. This is directly related to secondary analysis as an active and self-directed learning method. Allowing students to guide the content and goals of the research project teaches self-directed learning, which is important to enable students to later engage in lifelong learning (Blumberg 2005). In particular, inquiry-based learning increases students' scientific literacy and self-efficacy (Brickman et al. 2009).

Once students learn the methods and procedures of scientific inquiry, they will feel confident in applying this knowledge to later ventures. Students engaged in secondary analysis learn how to summarize data, make comparisons, test assumptions, and model factors associated with any number of outcomes. Most importantly, students learn how to identify and resolve problems within secondary datasets—skills that are incredibly valuable for future work as independent investigators or as members of a research team in any data-based profession.

A fourth advantage of the secondary analysis model is close *student-faculty collaboration*, which is associated with student satisfaction and retention (Nagda et al. 1998; Shellito et al. 2001). The research topic is initiated by students' interests, not faculty interests, although faculty advisors need to have adequate knowledge of the topics. Meetings typically take place in the faculty advisor's office, allowing the student and faculty member to work together through drafts of the research proposal, process feedback, and develop the products resulting at each step in the process. This close collaboration encourages student motivation and ownership of the research agenda, improving the likelihood of future publication.

Challenges of the Secondary-analysis Model

The secondary-analysis model will not always be effective. It requires dedication from both the student and faculty advisor. Students may become frustrated by the limitations of existing data and decide to frame a research question around available data instead of first developing the research question they want to answer (Smith et al. 2011). They may become frustrated by the many rules and regulations of statistical analysis, or have difficulty working with statistical programs on their own. If proper statistical software is not available through the institution, or if relevant data are not available free of charge, obtaining funding for these essential research components may be a challenge.

Faculty workloads may become burdensome if advisors mentor too many students or if one (or more) of their mentees requires extensive guidance. Departments may find that they do not have enough faculty members with appropriate experience and analytic skills (Hauhart and Grahe 2012). That is why it is important for the students and faculty advisors to meet before the semester begins to discuss student motivations, expectations, skill levels, and mutual interests.

Application to Health Sciences

The secondary analysis of an existing data model is especially effective in social science capstones for undergraduates pursuing careers in the health sciences. In addition to the general social science datasets mentioned above, several other publicly available health-related datasets are presented in Table 3. These datasets allow students to explore a health outcome that is important to their future professions using different disciplinary perspectives, including sociology, psychology, and anthropology. Some datasets contain additional health-relevant data on genetics, health assessments, laboratory values, and healthcare utilization.

The types of research produced by secondary analysis that are relevant to the health sciences can be classified as social epidemiology (e.g., why are certain groups of children more prone to obesity than others?) or health-services research (e.g., how is health literacy associated with treatment adherence among adults with chronic illness?). For example, a pharmacy student interested in the expansion of flu vaccination to pharmacy and other non-traditional settings analyzed the National H1N1 Flu Survey to identify patterns in vaccination settings based on socio-demographic factors (age, race, gender, income, etc.), insurance coverage, and beliefs about vaccinations. In another example, a pre-med student also interested in flu vaccination analyzed the same dataset to identify social patterns in parents' reasons for not vaccinating their children, again focusing on socio-demographic factors, health-care factors, and health beliefs. Both students have presented their research within the university and are preparing manuscripts for submission to peer-reviewed journals. A third student was less successful and decided not to pursue dissemination of the research after the capstone. These students' preliminary work with the secondary-research model informed expansion and revision of the model, which can now be assessed using rates of students' publication success, as well as the results of before-capstone and after-capstone testing of students' scientific literacy and research self-efficacy (Brickman et al. 2009).

Research experience is an important supplement to traditional clinical and scientific training for students interested in the

Table 3. Key Datasets for Social Science Research, By Disciplines

Dataset	Content	Relevant disciplines
General Social Survey[1]	Self-reported socio-demographics, behaviors, and attitudes of adults in the U.S.	Economics, political science, psychology, sociology
World Values Survey	Self-reported socio-demographics, social and political values, and well-being of adults in more than 80 countries	Anthropology, economics, political science, psychology, sociology
National Flu Survey[2,3]	Self-reported socio-demographics and beliefs, attitudes, and behaviors related to flu vaccination, data from adults and parents of children in the U.S.	Health-services research, medicine, pharmacy, public health
National Health Interview Survey[2]	Self-reported socio-demographics and health (status, care utilization, and behaviors), data from adults and parents of children in the U.S.	Health-services research, medicine, public health
National Health and Nutrition Examination Survey[2]	Self-reported socio-demographics, diet, health (status, behaviors, care utilization), and medication usage, data from adults and parents of children in the U.S.; data from physical examinations, including body measurement, blood tests, and dental screening.	Health-services research, medicine, nutrition, public health, pharmacy, psychiatry
Behavioral Risk Factor Surveillance System[2]	Self-reported socio-demographics and health (status and risk behaviors), data from adults and parents of children in each state in the U.S.	Medicine, public health, psychology
Chinese Longitudinal Healthy Longevity Survey[4]	Self-reported socio-demographics, attitudes, and health (status and behaviors) of centenarians, older adults, and middle-aged adults in China; DNA samples	Anthropology, demography, genetics, medicine, public health, sociology
Health and Retirement Survey[5]	Self-reported socio-demographics, labor-force participation, and health data from adults aged 50 and older in the U.S.; DNA samples	Economics, genetics, medicine, public health, sociology
National Survey of Children's Health[6]	Self-reported socio-demographics, health status, access to care, and social context for families of children in the U.S.	Medicine, pharmacy, psychology, public health

[1]Administered by National Opinion Research Center (NORC) at the University of Chicago
[2]Administered by the Centers for Disease Control and Prevention (CDC)
[3]The National H1N1 Flu Survey is an alternate dataset that is more easily available; it contains data relevant to both H1N1 and seasonal flu illness and vaccination
[4]Administered by Duke University, Peking University, and China Mainland Information Group
[5]Administered by the University of Michigan
[6]Administered by the Data Resource Center for Child and Adolescent Health. The related National Survey of Children with Special Health Care Needs focuses on households with children with special needs and is particularly relevant for occupational therapy.

Secondary Analysis of Existing Data continued

health professions. Secondary analysis of large national surveys allows students to examine health outcomes, risk factors, and the health-care system as experienced by thousands of people. This research model encourages students to think beyond clinical elements to understand how individual, interpersonal, and community-level factors all interact to affect the health of individuals and populations (Bronfenbrenner 1977).

Finally, social science research is a useful training experience for students who will take a revised version of the Medical College Admission Test in 2015 (https://www.aamc.org/students/download/266006/data/2015previewguide.pdf). Secondary analysis will teach skills in scientific inquiry and reasoning, including data-based statistical reasoning and reasoning about the design and execution of research. Content of the research projects will also reinforce new foundational concepts tested in the 2015 MCAT, including culture, inequality, population processes, and health disparities. Students who engage in secondary analysis of existing data early in their careers will receive this additional preparation for graduate-level entrance exams, and they also will have the opportunity to add a first-authored, peer-reviewed journal article to their resumes when they apply to professional schools or for employment.

Conclusion

The model of using secondary analysis of existing datasets emphasizes problem-solving and hypothesis-testing through statistical research in an individualized learning environment. These skills and experiences are vital to many types of careers that utilize statistical analysis, including research and program evaluation in social sciences, business, and medicine.

References

Beckman, Mary, and Nancy Hensel. 2009. "Making Explicit the Implicit: Defining Undergraduate Research." CUR Quarterly 29 (4): 40-44.

Bledsoe, Caroline H., Bruce Sherin, Adam G. Galinsky, Nathalia M. Headley, Carol A. Heimer, Erik Kjedgaard, James Lindgren, Jon D. Miller, Michael E. Roloff, and David H. Uttal. 2007. "Regulating Creativity: Research and Survival in the IRB Iron Cage." *Northwestern University Law Review* 101(2): 593-641.

Blumberg, Phyllis. 2005. "Why Self-directed Learning is not Learned and Practiced in Veterinary Education." *Journal of Veterinary Medical Education* 32(3): 290-295.

Brickman, Peggy, Cara Gormally, Norris Armstrong, and Brittan Hallar. 2009. "Effects of Inquiry-based Learning on Students' Science Literacy Skills and Confidence." *International Journal for the Scholarship of Teaching and Learning* 3(2). Retrieved from http://academics.georgiasouthern.edu/ijsotl/v3n2.html

Bronfenbrenner, Urie. 1977. "Toward an Experimental Ecology of Human Development." *American Psychologist* 32(7): 513-531.

Craney, Chris, Tara McKay, April Mazzeo, Janet Morris, Cheryl Prigodich, and Robert de Groot. 2011. "Cross-discipline Perceptions of the Undergraduate Research Experience." *The Journal of Higher Education* 82(1): 92-113.

Hauhart, Robert C, and Jon E. Grahe. 2010. "The Undergraduate Capstone Course in the Social Sciences: Results from a Regional Survey." *Teaching Sociology* 38(1): 4-17.

Hauhart, Robert C., and Jon E. Grahe. 2012. "A National Survey of American Higher Education Capstone Practices in Sociology and Psychology." *Teaching Sociology* 40(3): 227-241.

Lopatto, David. 2003. "The Essential Features of Undergraduate Research." *CUR Quarterly* 23(1): 139-142.

McKinney, Kathleen, and Melissa Busher. 2011. "The Sociology Research Experience Capstone Course at Three Institutions." *Teaching Sociology* 39(3): 290-302.

Nagda, Biren A., Sandra R. Gregerman, John Jonides, William von Hippel, and Jennifer S. Lerner. 1998. "Undergraduate Student-faculty Research Partnerships Affect Student Retention." *Review of Higher Education* 22: 55-72.

Puglisi, Tom. 2001. "IRB Review: It Helps to Know the Regulatory Framework." *Observer* 14(5). aps.psychologicalscience.org/index.php/uncategorized/irb-review-it-helps-to-know-the-regulatory-frameowrk.html.

Shellito, Cindy, Kalyn Shea, Gary Weissmann, Anke Mueller-Solger, and William Davis. 2001. "Successful Mentoring of Undergraduate Researchers: Tips for Creating Positive Student Research Experiences." *Journal of College Science Teaching* 30(7): 460-465.

Smith, Alexander K., John Z. Ayanian, Kenneth E. Covinsky, Bruce E. Landon, Ellen P. McCarthy, Christina C. Wee, and Michael A. Steinman. 2011. "Conducting High-Value Secondary Dataset Analysis: An Introductory Guide and Resources." *Journal of General Internal Medicine* 26(8): 920-929.

Trzesniewski, Kali H., M. Brent Donnellan, and Richard E. Lucas, eds. 2010. *Secondary Data Analysis: An Introduction for Psychologists*. American Psychological Association.

Jessica M. Sautter

U*niversity of the Sciences, j.sautter@usciences.edu*

Jessica Sautter is an assistant professor of medical sociology at the University of the Sciences in Philadelphia, Pennsylvania. She previously worked as a postdoctoral scholar in health-services research at the Duke University Medical Center and the Durham Veterans Administration Medical Center. Her research includes analysis of observational clinical studies and large secondary datasets, all focusing on social factors related to the health and well-being of children, adults, and older adults in the U.S. and China. Sautter advises undergraduates pursuing research capstones in sociology and social science, as well as master's theses in psychology.