

Data Citation at ICPSR

Jared Lyle, Elizabeth Moss, Christin Cave

Data Citation Workshop: Developing Policy & Practice

Washington, D.C.

12 July 2016

A partner in social science research

Find & Analyze Data

quick search for data

Membership in ICPSR

Deposit Data

ICPSR Summer Program

Resources for Instructors

Data Management & Curation

[About](#) | [Contact](#) | [Privacy](#) | [Careers](#)

ICPSR advances and expands social and behavioral research, acting as a global leader in data stewardship and providing rich data resources and responsive educational opportunities for present and future generations.



Announcements

[Webinar on Social Science Data Management and Curation Features](#)
[ICPSR Presenters](#)

[New Ten-Year Restricted-use NSDUH Data Released with Counties and Other Geographic Variables Added](#)

[ICPSR's 2014 Student Research Paper Deadline Approaching](#)

[More...](#)

Website Search



Archive Partners



MEASURES OF
EFFECTIVE TEACHING
Longitudinal Database

NACDA
Program on Aging

NATIONAL ARCHIVE
OF CRIMINAL
JUSTICE DATA



Child Care & Early Education
RESEARCH CONNECTIONS



SAMHDA

[1] What are we doing with data citation?

[2] How could better data citation impact us?

What are we doing with data citation?

ICPSR has provided data citations since 1990, and assigned digital object identifiers (DOIs) since 2008.

Study Description

Citation

Schneider, Barbara, and Linda J Waite. The 500 Family Study [1998-2000: United States]. ICPSR04549-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2008-05-30. <http://doi.org/10.3886/ICPSR04549.v1>

Persistent URL: <http://doi.org/10.3886/ICPSR04549.v1>

Export Citation:

- [RIS](#) (generic format for RefWorks, EndNote, etc.)
- [EndNote XML](#) (EndNote X4.0.1 or higher)

Machine-readable Data Citation

```
<div itemscope itemtype="http://schema.org/Dataset">
<h1 id="info" itemprop="name">The 500 Family Study [1998-2000: United States]
(ICPSR 4549) </h1>
<span itemprop="author" itemscope
itemtype="http://schema.org/Person"><span
itemprop="name">Schneider, Barbara</span>, <span
itemprop="affiliation">University of Chicago. National Opinion
Research Center (NORC). Alfred P. Sloan Center on Parents,
Children and Work</span>; </span>
<span itemprop="author" itemscope
itemtype="http://schema.org/Person"><span itemprop="name">Waite,
Linda J</span>, <span itemprop="affiliation">University of
Chicago. National Opinion Research Center (NORC). Alfred P.
Sloan Center on Parents, Children and Work</span></span>
<span itemprop="url">http://doi.org/10.3886/ICPSR04549.v1</span>
<span itemprop="datePublished">2008-05-30</span>
</div>
```

Attribution Requirement in Terms of Use

Citing Data

You agree to reference the recommended bibliographic citation in any publication that employs resources provided by ICPSR. Authors of publications based on ICPSR data are required to send citations of their published works to ICPSR for inclusion in a database of related publications (bibliography@icpsr.umich.edu✉) .

References

Bachman, J. G., Johnston, L. D., & O'Malley, P. M. (1998). Explaining recent increases in students' marijuana use: Impacts of perceived risks and disapproval, 1976 through 1996. *American Journal of Public Health, 88*, 887–892. doi:10.2105/AJPH.88.6.887

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Brown, T. L., Parks, G. S., Zimmerman, R. S., & Phillips, C. M. (2001). The role of religion in predicting adolescent alcohol use and problem drinking. *Journal of Studies on Alcohol, 62*, 696–705. Retrieved from <http://www.csa.com>

Caviness, L. L., Cofey, L. E., Cunningham, D. B., Griffin, A. B., Liu, S., Martin, P., . . . Zelko, H. E. (2007). *2006 National Survey on Drug Use and Health: Data collection final report*. Rockville, MD: Substance

Schreiber, J. B., Stage, F. K., King, J., Nora, A., & Barlow, E. A. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research, 99*, 323–337. doi:10.3200/JOER.99.6.323-338

Stern, M. K., & Wiens, B. A. (2009). Ethnic differences in adolescent perceptions of and attitudes toward substance use. *Journal of Ethnicity in Substance Abuse, 8*, 54–69. doi:10.1080/15332640802683417

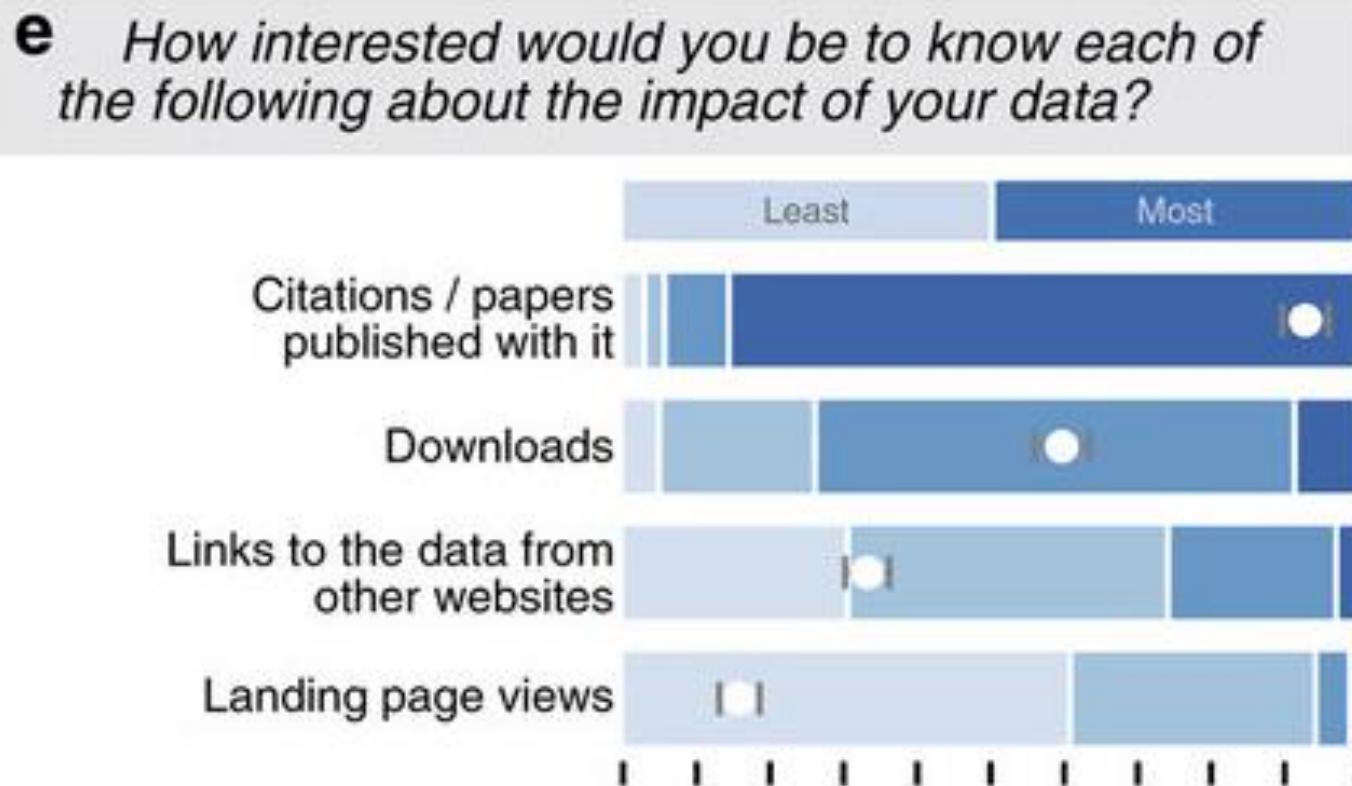
U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies. (2009a). National Survey on Drug Use and Health, 2006 [Computer file]. ICPSR21240-v4. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-08-12. doi:10.3886/ICPSR21240

U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

[doi:10.3886/ICPSR21240](https://doi.org/10.3886/ICPSR21240)

Our users value data citations

Researchers (n=247) were asked:
“How interested would you be to know each of the following about the impact of your data?



*White dots show the mean on a scale of one-to-four. All error bars depict 95% confidence intervals calculated by basic bootstrap with 10,000 resamplings.

- **Downloads:** Download counts, on the other hand, are both highly valuable and practical to collect. Downloads were a resounding second-choice metric for researchers and 85% of repositories already track them.
- **Citations:** Citations are the coin of the academic realm. They were by far the most interesting metric to both researchers and data managers. Unfortunately, citations are much more difficult than download counts to work with, and relatively few repositories track them. Beyond technical complexity, the biggest challenge is cultural: data citation practices are inconsistent at best, and formal data citation is rare. Despite the difficulty, the value of citations is too high to ignore, even in the short term.

Funders also value data citations



Student-Athlete Experiences
Data Archive



Child Care & Early Education
RESEARCH CONNECTIONS



Data citation allows us to answer:

- Who uses the data?
- How are they used?
- With what impact?

Track and Link Data and Publications



Find & Analyze Data

Find Data

Search/Compare Variables

Find Publications

Resources for Students

Get Help

About the Bibliography

The *ICPSR Bibliography of Data-related Literature* is a searchable database that as of 2016 contains over 70,000 citations of published and unpublished works resulting from analyses of data held in the ICPSR archive. Developed with support from the National Science Foundation (SES-9977984), the *Bibliography* represents over 50 years of scholarship in the quantitative social sciences, extending from the inception of ICPSR in 1962 to the present. [\(more\)](#)

Find Publications

The *ICPSR Bibliography of Data-related Literature* is a continuously-updated database of thousands of citations of [works using data held in the ICPSR archive](#). The works include journal articles, books, book chapters, government and agency reports, working papers, dissertations, conference papers, meeting presentations, unpublished manuscripts, magazine and newspaper articles, and audiovisual materials.

- [View all citations](#)
- [Browse by author](#)
- [Browse by journal](#)
- [Browse by study](#)

Search Tips

- Our citation search indexes only the literal citations, not the full text of the publications.
- Instead of entering a research question, use only one or two search words, and use the filters to narrow

Filter Results

Search Results

Pub. Year ▾

Pub. Type ▾

Journal ▾

Author ▾

Study ▾

Export Results ▾

Use the search box below to revise your search, or [start a new search](#).

 REVISE SEARCH

Studies

Variables

Publications (69,384)

Your query returned 69384 citations.

Sort by: Pub Date (newest) ▾

- Pub Date (newest)
- Title
- Recent Additions
- Author

Export

Clear All

Pub.
Year

Citation

1.	<input type="checkbox"/>	2016	DeLisi, Matt, Piquero, Alex R., Cardwell, Stephanie M. The unpredictability of murder: Juvenile homicide in the Pathways to Desistance Study . <i>Youth Violence and Juvenile Justice</i> . 14, (1), 26-42. Full Text Options: DOI ↗ Worldcat ↗ Google Scholar ↗ Export Options: RIS EndNote XML
2.	<input type="checkbox"/>	2016	Helms, Ronald, Gutierrez, Ricky S., Reeves-Gutierrez, Debra . Jail mental health resourcing: A conceptual and empirical study of social determinants . <i>International Journal of Offender Therapy and Comparative Criminology</i> . 60, (9), 1036-1063. Full Text Options: DOI ↗ Worldcat ↗ Google Scholar ↗ Export Options: RIS EndNote XML
3.	<input type="checkbox"/>	2016	Palamar, Joseph J., Keyes, Katherine, Cleland, Charles M. Underreporting of ecstasy use among high school seniors in the US . <i>Drug and Alcohol Dependence</i> . Full Text Options: DOI ↗ Worldcat ↗ Google Scholar ↗ Export Options: RIS EndNote XML

Related Publications (?)

- [List all ~199 citations associated with this study](#)
- [Search the citations in this study](#)

Most Recent Publications

2014	Rycraft Nicholson, Cara . Predictors of Adolescent Delinquency: Comparing Models and Additive Effects . Dissertation, Roosevelt University. Export Options: RIS/EndNote
2013	Erol, R.Y., Orth, U. Actor and partner effects of self-esteem on relationship satisfaction and the mediating role of secure attachment between the partners . <i>Journal of Research in Personality</i> . 47, (1), 26-35. Full Text Options: DOI ➔ Worldcat ➔ Google Scholar ➔ Export Options: RIS/EndNote
2013	Offer, Shira . Family time activities and adolescents' emotional well-being . <i>Journal of Marriage and Family</i> . 75, (1), 26-41. Full Text Options: DOI ➔ Worldcat ➔ Google Scholar ➔ Export Options: RIS/EndNote
2011	Offer, Shira, Schneider, Barbara . Revisiting the gender gap in time-use patterns: Multitasking and well-being among mothers and fathers in dual-earner families . <i>American Sociological Review</i> . 76, (6), 809-833. Full Text Options: DOI ➔ Worldcat ➔ Google Scholar ➔ Export Options: RIS/EndNote
2011	Weston, William . The college class at work and home . <i>Society</i> . 48, (3), 236-241. Full Text Options: DOI ➔ Worldcat ➔ Google Scholar ➔ Export Options: RIS/EndNote
2009	Inter-university Consortium for Political and Social Research . Exploring the Second Shift: A Data-Driven Learning Guide . Inter-university Consortium for Political and Social Research. Full Text Options: DOI ➔ Export Options: RIS/EndNote

Citation

The following citation is from the [ICPSR Bibliography of Data-related Literature](#):

Offer, Shira, Schneider, Barbara . Dec 2011. **Revisiting the gender gap in time-use patterns: Multitasking and well-being among mothers and fathers in dual-earner families.** *American Sociological Review*. 76, (6), 809-833. doi:10.1177/0003122411425170

Full Text Options: [DOI](#) ↗ [WorldCat](#) ↗ [Google Scholar](#) ↗

Export Citation: [RIS](#) [EndNote](#) [XML](#)

Related Studies

This publication is related to the following dataset(s):

- [The 500 Family Study \[1998-2000: United States\]](#)

[Download PDF](#)[Export](#) [Advanced search](#)

Social Science Research

Volume 53, September 2015, Pages 288–299



Seclusion, decision-making power, and gender disparities in adult health: Examining hypertension in India

Samuel Stroope  

[Show more](#)

doi:10.1016/j.ssresearch.2015.05.013

[Get rights and content](#)

Highlights

- Women's seclusion is associated with higher odds of women's hypertension.
- Women's seclusion is associated with lower odds of men's hypertension.
- Women's low decision-making power is associated with higher odds of women's hypertension.

▼ Recommended articles

[The role of provider supply and organization...](#)

2013, Social Science & Medicine [more](#)

[Scientific consensus, the law, and same sex...](#)

2015, Social Science Research [more](#)

[The association of daily physical symptoms ...](#)

2015, Social Science & Medicine [more](#)

[View more articles »](#)

► Citing articles (1)

► Related book content

[Data for this Article](#)

 ICPSR

Interuniversity Consortium for
Political and Social Research
Social and behavioral science research
data

 A red oval highlights the ICPSR logo and text.

All Databases

Select a Database

Data Citation Index

Additional Resources

Search

Cited Reference Search

Advanced Search

Search History

Data Citation IndexSM

<< Back to results list

Record 1 of 3

Record from Data Citation Index SM



(0)



Save to:

ENDNOTE[®] WEBENDNOTE[®]

RefWorks

I Wrote These Publications



more options

Chinese Household Income Project, 1988.

From Repository: Inter-university Consortium for Political and Social Research.

Author(s): Griffin, Keith; Renwei, Zhao

Source: Inter-university Consortium for Political and Social Research DOI: <http://dx.doi.org/10.3886/ICPSR09836> (Viewed Date: 09 Aug 2012) Published Year: 2010 Version: v2

Cited References: 0

Abstract: The purpose of this project was to measure and estimate the distribution of income in both rural and urban areas of the People's Republic of China. The principal investigators based their definition of income on cash payments and on a broad range of additional components: payments in kind valued at market prices, agricultural output produced for self-consumption valued at market prices, the value of ration coupons and other direct subsidies, and the imputed value of housing. The rural component of this collection consists of two datafiles, one in which the individual is the unit of analysis and a second in which the household is the unit of analysis. Individual rural respondents reported on their employment status, level of education, Communist Party membership, type of employer (e.g., public, private, or foreign), type of economic sector in which employed, occupation, whether they held a second job, retirement status, monthly pension, monthly wage, and other sources of income. Demographic variables include relationship to household, gender, age, and student status. Rural households reported extensively on the character of the household and residence. Information was elicited on type of terrain surrounding the house, geographic position, type of house, and availability of electricity. Also reported were sources of household income (e.g., farming, industry, government, rents, and interest), taxes paid, value of farm, total amount and type of cultivated land, financial assets and debts, quantity and value of various crops (e.g., grains, cotton, flax, sugar, tobacco, fruits and vegetables, tea, seeds, nuts, lumber, livestock and poultry, eggs, fish and shrimp, wool, honey, and silkworm cocoons), amount of grain purchased or provided by a collective, use of chemical fertilizers, gasoline, and oil, quantity and value of agricultural machinery, and all household expenditures (e.g., food, fuel, medicine, education, transportation, and electricity). The urban component of this collection also consists of two data files, one in which the individual is the unit of analysis and a second in which the household is the unit of analysis. Individual urban respondents reported on their economic status within the household, Communist Party membership, sex, age, nature of employment, and relationship to the household head. Information was collected on all types and sources of income from each member of the household whether working, nonworking, or retired, all revenue received by owners of private or individual enterprises, and all in-kind payments (e.g., food and durable and non-durable goods). Urban households reported total income (including salaries, interest on savings and bonds, dividends, rent, leases, alimony, gifts, and boarding fees), all types and values of foodstuffs received, and total debt. Information was also gathered on household accommodations and living conditions, including number of rooms, total living area in square meters, availability and cost of running water, sanitary facilities, heating and air-conditioning equipment, kitchen availability, location of residence, ownership of home, and availability of electricity and telephone. Households reported on all of their expenditures including amounts spent on food items such as wheat, rice, edible oils, pork, beef and mutton, poultry, fish and seafood, sugar, and vegetables by means of both coupons from state-owned stores and at free market prices. Information was also collected on rents paid by the households, fuel available, type of transportation used, and availability and use of medical and childcare.

Document Type: Data study

Data Type: survey data

Accession Number: DRCI:DATA2012175001845994

Language: English

Funding:

Funding Agency	Grant Number
Ford Foundation	

Times Cited: 9

Create Citation Alert

This article has been cited 9 times in Web of Knowledge.

Hannum, E. Market transition, educational disparities, and family strategies in rural China: New evidence on gender stratification and development. *DEMOGRAPHY*, MAY 2005.

Liu, ZQ. The economic impact and determinants of investment in human and political capital in China. *ECONOMIC DEVELOPMENT AND CULTURAL CHANGE*, JUL 2003.

Shu, XL. Market transition and gender gap in earnings in urban China. *SOCIAL FORCES*, JUN 2003.

[view all 9 citing articles]

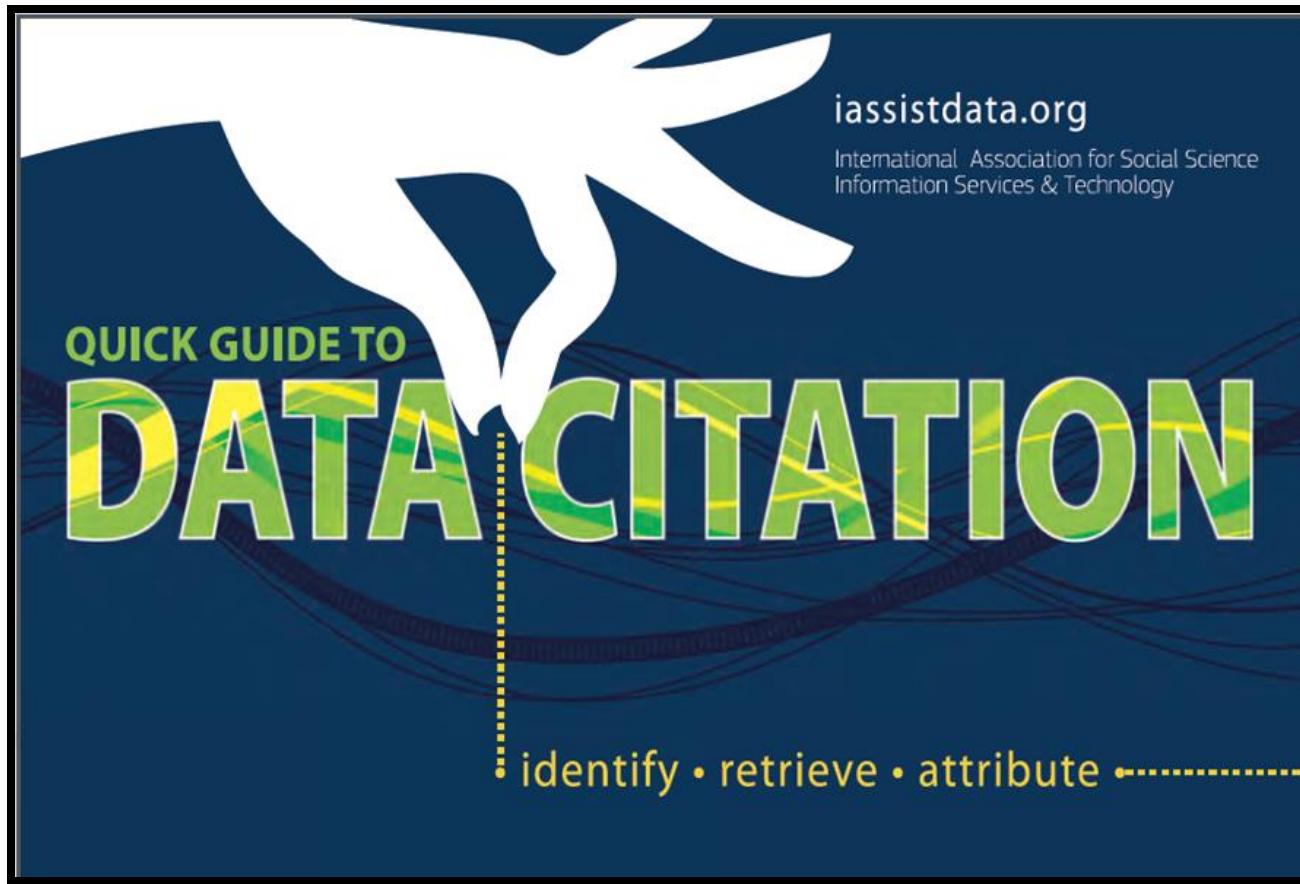
Cited References: 0

Additional Information

How to cite this Resource

Suggest a correction

If you would like to improve the quality of the data in this record, please [suggest a correction](#).



<http://www.iassistedata.org/blog/iassist-publishes-quick-guide-data-citation>

How could better data citation impact us?

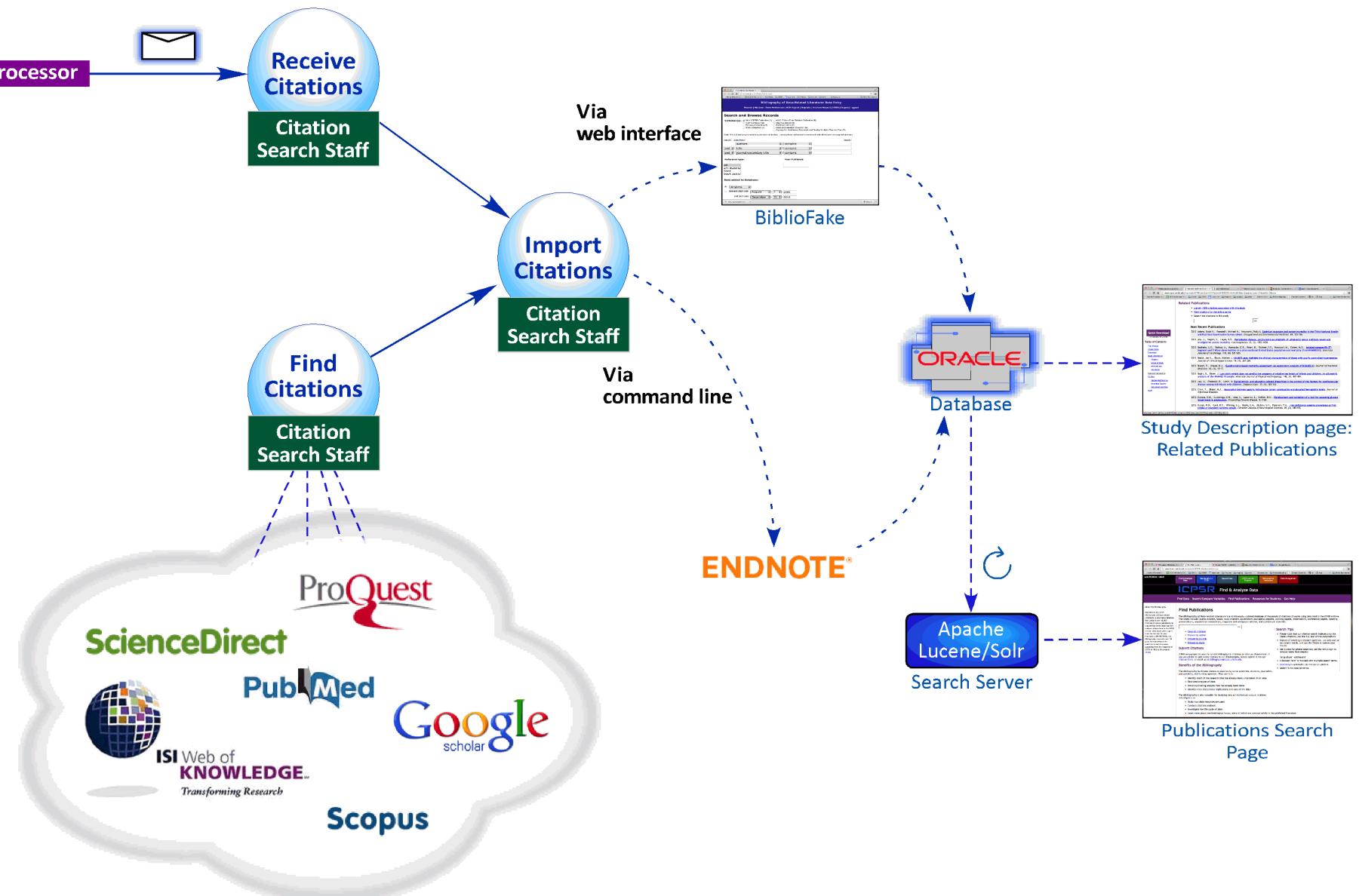
Bibliography

My own mind was my source
due to experience.

<http://www.flickr.com/photos/papertrix/38028138/>

(CC BY-NC 2.0)

Populating the ICPSR Bibliography



Challenges of Data Citation

- Poor and inconsistent citing practices
- Emerging data citation standard
- Ambiguous descriptions of data used in abstract, methodology, acknowledgments
- Requires inefficient (human) searching and browsing to track data and keep up with the demand
- Without standard practice, it is very difficult to quantify the impact of data sharing

Sample?
Methods?



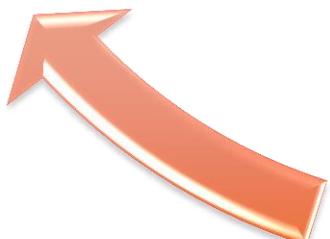
Abstract?
Acknowledgements?



**Data
“Sighting”
(implicit)
vs.
Data
Citing
(explicit)**



Discussion?
Footnotes?



Charts and
Tables?



Appendices?
References!

Examples of poor citation practice

- Sample described, not named, no author information, no access information, only a publication cited
- Data named in text, with some attribution, but no access information
- Cited in reference section, but with no permanent, unique identifier, so difficult for indexing scripts to find to automate tracking

Examples of a poor data citation

Poorly described and cited data

+ differently. Using survey data from 272 staff who worked at a Midwestern maximum-security state prison, this study examined the association between the occupational stressors of perceived dangerousness of the job, role conflict, role ambiguity, repetitiveness, and work-on-family conflict with the three forms of organizational commitment. The effects of the occupational stressors varied for each form of organizational commitment. Specifically, work-on-family conflict had a significant positive association with continuance commitment. Role conflict and repetitiveness had negative associations with moral commitment. Finally, all five stressors had significant negative associations with affective commitment.

Excessive human search effort, extensive collection knowledge

=

Too costly, too questionable for confident measure of impact

Examples of a good data citation

Citing data with a DOI

+

Monto, M. A. (1999a). *Clients of street prostitutes, in Portland, Oregon, San Francisco and Santa Clara, California, and Las Vegas, Nevada, 1996-1999* (ICPSR version). Portland, OR: University of Portland [Producer], Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributor], 2000. doi:10.3886/ICPSR02859.v1

Smith, T. W., Marsden, P. V., & Hout, M. *General Social Survey, 1972-2010* [Cumulative File] (ICPSR31521-v1). Storrs: Roper Center for Public Opinion Research, University of Connecticut/Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributors], 2011-08-01. doi:10.3886/ICPSR31521.v1

ICPSR unique
DOI

ICPSR Study
Number

Version of the data file

Minimal human search effort

=

High hit accuracy for the cost, and better confidence of impact measures



Google Scholar Alerts <sch... Mar 29 (4 days ago) 

to me 



Scholar Alert: ["doi:10.3886"]; Articles excluding patents

[Ordinary or Peculiar Men? Comparing the Customers of Prostitutes With a Nationally Representative Sample of Men](#)

MA Monto, C Milrod - [International Journal of Offender Therapy and](#) ... , 2013

Page 1. [International Journal of Offender Therapy and Comparative Criminology](#) XX(X) 1 –19 ©

The Author(s) 2013 Reprints and permissions: [sagepub.com/journalsPermissions.nav](#) DOI:

10.1177/0306624X13480487 [ijo.sagepub.com](#) Article Ordinary or Peculiar Men? ...

This Google Scholar Alert is brought to you by Google.

[Cancel alert](#)

Make Your Data Count!

- If it's not cited, it can't be counted
- Without counting data use, there is no accurate way to measure the impact of your shared data
- Without a well-formed citation, your data cannot take advantage of the potential of linked scholarly publishing
- Store your data where citations are unique and persistent
- Cite your own data and others' in your publications

Thank you!

Jared Lyle
lyle@umich.edu

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

11. *Leucosia* (Leucosia) *leucostoma* (Fabricius) (Fig. 11)

11. *Leucosia* (Leucosia) *leucostoma* (Fabricius) (Fig. 11)



ISR

INSTITUTE FOR SOCIAL RESEARCH

THE UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN

1 2 3 4 5 6 7 8 9 10 11
IBM U25096

Audience

Sep 24, 2012 - Sep 23, 2013



Unique Visitors

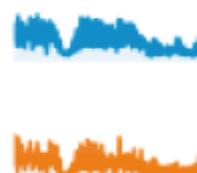
Studies
308,967



Bibliography
47,379



SSVD
37,753



Pageviews

Studies
4,563,388



Bibliography
898,195



SSVD
1,442,353

Source: Google Analytics



The Scholix initiative is a high level interoperability framework for exchanging information about the links between scholarly literature and data. It aims to build an open information ecosystem to understand systematically what data underpins literature and what literature references data.

<http://www.scholix.org/>

**Placeholder to request persistent IDs
for government reports**

No Common Practice of Formal Data Citation

- Abstract?
- Acknowledgements?
- Charts and Tables?
- Appendices?
- Discussion?
- Footnotes?
- Sample?
- Methods?
- **References!**
- Without an explicit citation, reader must infer or be out of luck
- No attribution—no credit
- No access—no reuse
- No discernible impact!

Example

Abstract A huge literature shows that teen mothers face a variety of detriments across the life course, including truncated educational attainment. To what extent is this association causal? The estimated effects of teen motherhood on schooling vary widely, ranging from no discernible difference to 2.6 fewer years among teen mothers. The magnitude of educational consequences is therefore uncertain, despite voluminous policy and prevention efforts that rest on the assumption of a negative and presumably causal effect. This study adjudicates between two potential sources of inconsistency in the literature—methodological differences or cohort differences—by using a single, high-quality data source: namely, The National Longitudinal Study of Adolescent Health. We replicate analyses across four different statistical strategies: ordinary least squares regression;

Data

We use data from Add Health (the National Longitudinal Study of Adolescent Health), a school-based, nationally representative sample of 20,745 seventh through twelfth graders in 1994–1995. Respondents were reinterviewed in 1996 (Wave II), 2001–2002 (Wave III),

Acknowledgments Kane is the corresponding author; she carried out most of the statistical analysis and wrote the first draft of this article. She was also responsible for coordinating input from coauthors on subsequent drafts. This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwistle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>). This research received support