Reading Between the Lines:
Using Citations to Understand Anthropologists' Reading Patterns

Madeleine Sturma and Celia Emmelhainz

University of California, Berkeley, 230 Kroeber Hall, Berkeley CA 94720, USA

Abstract: Academic libraries want to collect the materials most useful to researchers, yet how can libraries know how successful they are? While Berkeley's George and Mary Foster Anthropology Library collects data on which books circulate, it is difficult to evaluate how materials are actually being used to further the discipline of anthropology. In this article, we examine sources cited by our a) faculty members, b) dissertation writers, and c) honors thesis students to better understand how anthropologists read when conducting research. This paper compares materials used across subfields and research levels to highlight patterns in citations within this discipline, leading to new insights that will improve collection development among anthropology librarians.

1. Introduction

What are anthropologists currently citing, and does this vary by their level of experience? Do citation patterns vary by subfield--and how can anthropology librarians most effectively support researchers in these disciplines? This paper seeks to understand how anthropologists at the University of California, Berkeley cite books and articles from our collection, through a comparative citation analysis of undergraduate, graduate, and faculty publications and theses from the past ten years. The goal is for the Anthropology Library at UC Berkeley to improve its ability to support scholars in the midst of changing scholarly practices.

2. Literature Review

Prior researchers have used citation analyses to move beyond looking at individual research habits and into understanding broader patterns in scholarly use of academic library collections. Datig (2016) reviewed citations by honors

students in the social sciences at NYU Abu Dhabi, using focus groups and a citation analysis to understand the sources students used and where they found them. She concludes that professors have a heavy influence on student source choice, and that undergraduate students prefer journal articles.

At the graduate student level, Schadl and Todeschini (2015) examined how Latin American Studies dissertation writers at the University of New Mexico used Spanish and Portuguese sources, comparing citations to the university library collections to gauge whether the collections meets student needs. The authors find Spanish sources central to graduate research, yet find it challenging to assess how the library collection is actually being used by students.

At the University of Colorado Boulder, Kellsey and Knievel (2012) assessed faculty citations in comparison with how many sources the library owned, how outside sources were being acquired by researchers, and the level of interdisciplinary usage. They found that most sources cited were indeed owned by the research institution.

This project was inspired by several collaborative citation analyses led by Susan Edwards at UC Berkeley, including a study of how dissertation writers use non-English language materials (Edwards et al., 2017), which found that 7% of citations in anthropology dissertations between 2008-2015 were of non-English languages, particularly Spanish and French--and yet 345 of 398 languages collected by the Library had no citations within an eight-year period. In a follow up assessment of social science faculty, Edwards et al (2019) found 9% of anthropology faculty citations in 2013-2017 were of non-English sources, with French, Spanish, Russian, Chinese, and German emerging as key languages cited across the four departments analyzed (anthropology, history, political science, and sociology). Yet again, 373 of 398 languages collected by the Library were not cited by faculty in these four highly international departments in a five year period.

3. Methods

As a research institution, UC Berkeley provides a scholarly home for anthropologists at three levels of study: undergraduate, graduate, and faculty. In this department, researchers appear to cluster in three subfields: cultural anthropology, archaeology, and biological/medical anthropology (Emmelhainz and Estrada, 2018). To better understand citation habits across such research

levels and subfields, we extracted, cleaned, and merged three datasets from the years 2008-2018.

First, we used the Scopus API to extract 4,996 citations from 119 articles and chapters by Berkeley anthropologists published in 2013-2017. Prior research (Emmelhainz and Estrada 2018) suggests this is half the known output of 253 tenured or tenure-track faculty publications in this time period. We hope to expand our faculty data back to 2008, but that requires manual data entry and cleaning and was not ready in time for this analysis. The 2013-2017 data focuses on international researchers and excludes several US-based researchers (Edwards et al 2018).

Second, we cleaned an existing dataset of 27,260 citations from 87 anthropology dissertations openly deposited by Berkeley graduate students with ProQuest in 2008-2015; using commencement brochures, we estimate this is approximately half of the 150 anthropology dissertations written at Berkeley in this time period. (All dissertations were deposited with the Library and ProQuest, but may have been embargoed at ProQuest by author request, meaning metadata was unavailable for extraction for this project).

Third, we scanned 5,605 citations from all 114 undergraduate honors theses deposited at Berkeley's George and Mary Foster Anthropology Library in 2008-2018, representing half the 210 theses listed in commencement brochures (the other theses do not seem to be been deposited on campus).

We cleaned this data in Microsoft Excel and Open Refine, standardizing author names to first author, item title, and publisher names, estimating item type (book, article, chapter, news, web), and assigning a disciplinary 'subfield' to our faculty and undergraduate authors based on faculty profiles on the department website. This is approximate but allows for comparison of citation patterns between e.g. cultural anthropologists and archaeologists. For the purposes of comparison, we assume that faculty typically advise students within the same subfield.

After cleaning the data in Excel and Open Refine, we ended up with a dataset of 38,000 citations from 320 scholarly works completed within the department in the last ten years. This represents half the undergraduate theses from 2008-2018, half the doctoral dissertations from 2008-2015, and half the faculty publications

from 2013-2017. We recognize that which theses were deposited in the library, which dissertations were embargoed when deposited with ProQuest, and which faculty publications were indexed in Scopus may introduce bias into this analysis. We still found insight in this broad overview of citations within a department, allowing us to understand how anthropologists are deploying literature reviews at different levels.

Our resulting dataset includes the following variables for department publications: author type (faculty, graduate, undergraduate) subfield (cultural, biological, archaeology), author, and date. It includes the following variables for items cited: type (book, article), year, title, publisher, and language. We used Excel pivot tables and Tableau to visualize patterns and better understand which publishers, journals, key works, and item types are most used by scholars at different levels and in different subfields of anthropology, to create a public dataset that could be reused, and to lay the groundwork for future comparisons with our existing library collections.

4. Results & Discussion
Type of work cited

|	Faculty	Graduate	Undergraduate
Article	38.86%	31.16%	34.38%
Book	34.64%	44.50%	36.02%
Chapter	25.67%	19.41%	16.44%
News	0.79%	3.48%	4.62%
Web	0.04%	1.45%	8.54%

Faculty, graduate students, and undergraduates cite broadly similar types of material, with articles, books, chapters, news, and web articles emerging as top sources. Note that interviews and fieldwork are also key sources in anthropological research, but rarely listed individually in reference lists. When compiling a literature review, all scholars drew primarily on articles, full books, and chapters. Faculty cite articles most often, while graduate and undergraduate students relied more heavily on books. Undergraduate students used more news and web sources, which may function as a "primary source" for thesis writers who lack access to international research sites for conducting original research.

Year of work cited

Anthropologists most often cite works published in the late 1980s to 2010s, and scholars at all levels and in both cultural anthropology and archaeology cite works from broadly similar time periods.

Years of Sources Cited in Cultural Anthropology
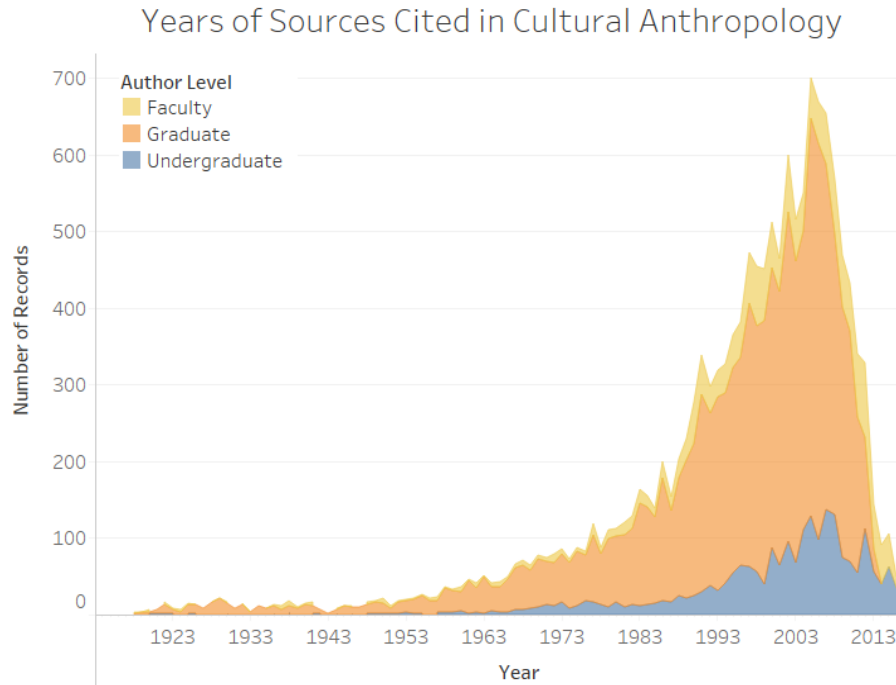
A long tail stretches back to works in the 1600s (not shown). Broadly comparable peaks and troughs suggest that publications prior to 1980, while valued classics, are no longer cited as heavily in anthropology.

Years of Sources Cited in Archaeology
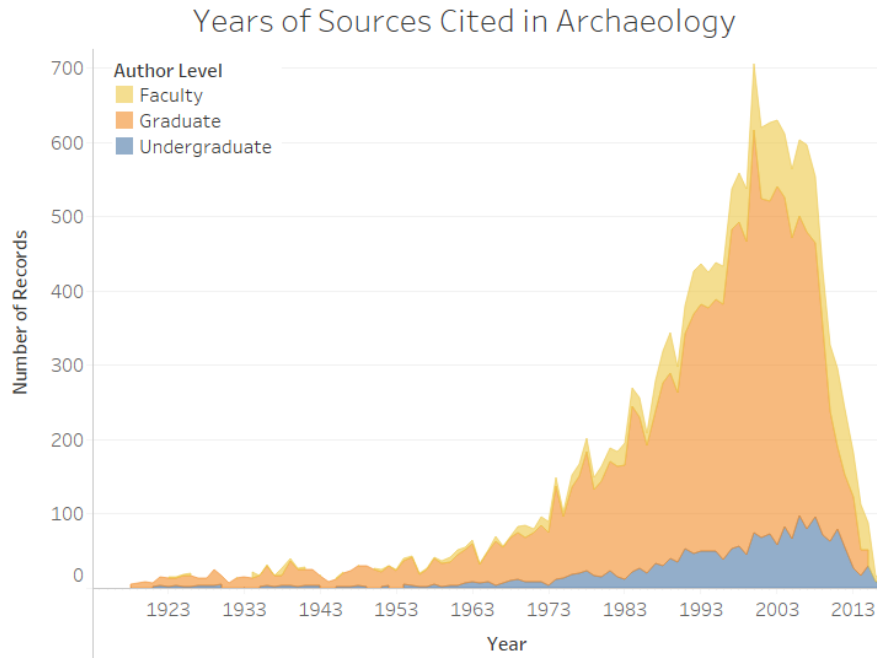
Languages Cited

While anthropologists conduct fieldwork in many languages, they cite literature primarily in English, representing 94% of the 35,000 citations. Only 5.5% of references were to works in other languages, which was most frequent at the faculty level (9%), moderate among graduate students (5% of all dissertation citations), and lowest among undergraduates (2% of all citations).

Spanish was the most cited language besides English, representing 2.5% of citations in this dataset, followed by French, Portuguese, Japanese, and Russian. The presence of less commonly taught languages such as Latvian and Chinese can often be attributed to a single person's research. This chart illustrates the tension between our desire as a library to collect broadly in all languages, and the low use of sources in languages other than English by most of our scholars.

Non-English Languages Most Cited
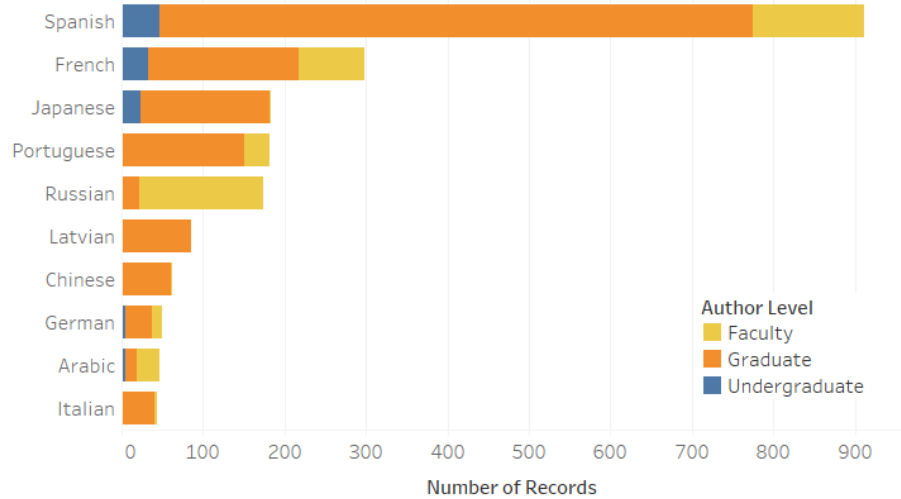
Books most cited

Imagined Communities	Anderson, Benedict	41
Outline of a Theory of Practice	Bourdieu, Pierre	35
Discipline and punish: the birth of the prison	Foucault, Michel	31
The Practice of Everyday Life	de Certeau, Michel	28
Anthropos Today: Reflections on Modern Equipment	Rabinow, Paul	24
The Logic of Practice	Bourdieu, Pierre	23
French Modern: Norms and Forms of the Social Environment	Rabinow, Paul	22
The History of Sexuality	Foucault, Michel	21
Modernity at Large: Cultural Dimensions of Globalization	Appadurai, Arjun	20
Orientalism	Said, Edward	19
The Gift	Mauss, Marcel	18
How to Do Things with Words	Austin, John L.	16
Publics and Counterpublics	Warner, Michael	15
A Thousand Plateaus: Capitalism and Schizophrenia	Deleuze, Gilles	14

Local anthropologists cite a shared body of works in anthropological philosophy; multiple titles from Pierre Bourdieu, Michel Foucault, and Paul Rabinow fall in the top 15 books most cited. All are male. A smaller sample of faculty and undergraduate works meant it was not useful to break out citations by level. With a larger sample, it would be interesting to compare trends over time.

Publishers most cited
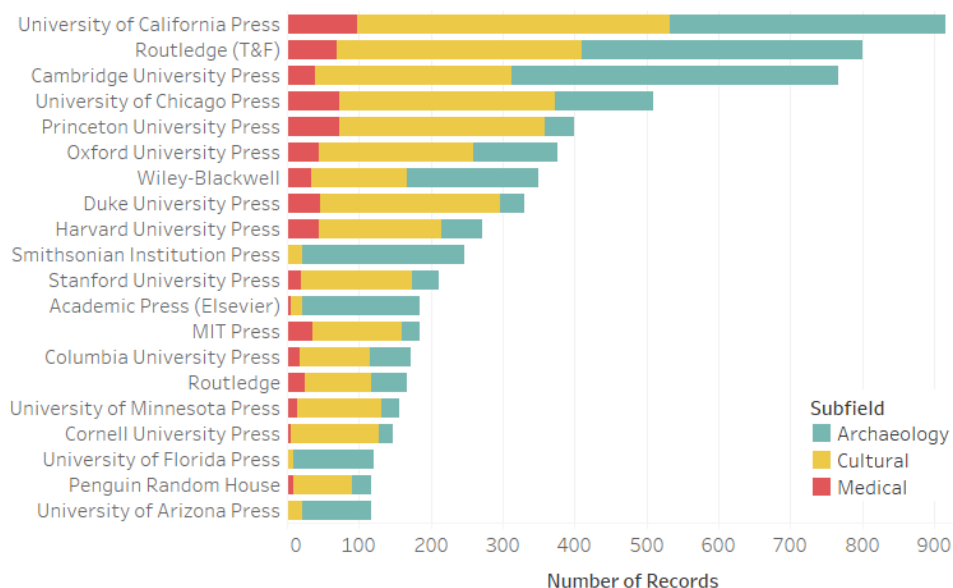
Key book publishers in anthropology vary by subfield, with citations in biological anthropology drawing from technical and scientific publishers. Medical and cultural anthropologists relied on university and private presses, while archaeologists cite university presses and some companies specializing in archaeological research. (Note that when cleaning our data, we merged variations on a publisher's name, and merged smaller or older imprints into the largest grouping, to highlight current consolidation in the market).

Often cited university presses include the University of California Press, Cambridge University Press, University of Chicago Press, and Princeton University Press. Among private publishers, Routledge monographs are frequently cited--a challenge for librarians given the high cost of acquisition. Graduate students cite university presses more often, while undergraduate and faculty authors cite more from private publishing houses.

Journals most cited

We looked first at the journals most cited, where *American Antiquity* emerges as a clear leader, then considered citations by subfield. Journals like the *American Journal of Physical Anthropology* are cited by biological anthropologists and archaeologists, while *Current Anthropology* is cited by cultural anthropologists

and archaeologists; other journals are read by a single specialty. This reinforces the need for librarians to attend to the specializations in the departments they serve, as there may not be a single core of journals for all researchers.

Top Journals by Subfield
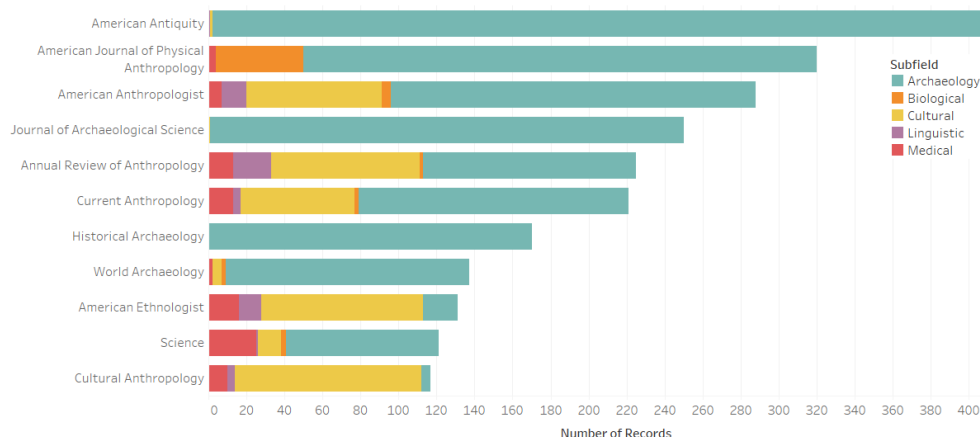
5. Conclusions and Further Study

As we compared citations within the anthropology department at UC Berkeley, we found strong differences across scientific, social science, and humanistic subfields, and relatively weak differences by academic level (undergraduate, graduate, and faculty). This suggests that when developing collections or providing research support, it is as important for librarians to learn about the specializations of their scholars as to learn about the developmental needs of scholars at different levels.

Becoming more familiar with key journals, titles, and publishers has helped us to refine our anthropology library's collection management and become more aware of trends in scholarly publishing, as we saw with the consolidation of academic publishers. In the future, we hope to cross-reference citations against our library holdings and talk with faculty and students to find out how they work around gaps in our library collection, something we have previously labeled "creative workarounds" (Emmelhainz, 2015). Another project would be to compare how often anthropologists cite within or across academic disciplines--and to assess whether anthropologists at historic universities such as Berkeley are more inclined to engage in self-citation within their home institution. Finally, any trends in citation over time are worth examining in more detail. We hope to release our existing dataset, allowing others to build on it in asking additional questions of their own.

References
Datig, Ilka. (2016). Citation Behavior of Advanced Undergraduate Students in the Social Sciences: A Mixed-Method Approach, *Behavioral & Social Sciences Librarian*, 35:2, 64-80, doi:10.1080/01639269.2016.1214559.

Edwards, Susan, Lynn Jones, and Scott Paul McGinnis. (2017). Big Data for Big Questions: Assessing the Impact of Non-English Language Sources on Doctoral Research at Berkeley. *ACRL Proceedings*, 13. Baltimore: American Library Association. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/2017/BigDataforBigQuestions.pdf

Edwards, Susan, Chan Li, Celia Emmelhainz, Adam Clemons, Liladhar Pendse, and Natalia Estrada. (2018). Collecting Globally, Connecting Locally: 21st Century Libraries. *Library Assessment Conference Proceedings*, 14. Retrieved from https://www.libraryassessment.org/wp-content/uploads/2019/09/76-Edwards-et-al-ConnectingGlobally.pdf

Emmelhainz, Celia. (2015). Supporting the expatriate social scientist: Faculty research and information access in Post-Soviet Kazakhstan. *Journal of Librarianship and Information Science*, 49:1, 69-81.

Emmelhainz, Celia, and Estrada, Natalia. (2018). Searching for recent anthropology and archaeology publications. *ACRL-ANSS Currents* 33:1, 17-21. escholarship.org/uc/item/94c8t9sk

Kellsey, Charlene, & Knievel, Jennifer (2012). Overlap between Humanities Faculty Citation and Library Monograph Collections, 2004–2009. *College & Research Libraries*, 73:6, 569-583, doi.org/10.5860/crl-280.

Schadl, Suzanne M., and Marina Todeschini (2015). Cite Globally, Analyze Locally: Citation Analysis from a Local Latin American Studies Perspective *College & Research Libraries*, 76:2, 136-149, doi:10.5860/crl.76.2.136.