

Efficiency in Scholarship: Do Keywords Matter?

Andrew Abbott
University of Chicago

I Scholars and Libraries.

Current debates in the library world often involve claims about the power or the danger of new research tools. On the one hand, they are said to make research possible that has never been seen before. On the other, they are said to automate or undercut the human factor in scholarship. Both sides make explicit or implicit claims about the relative quality of scholarship before and after the digital revolution.

Such claims presuppose a measure of scholarly quality to be optimized or improved. Yet even in the loosest sense, there is at present neither a serious measure nor even a real concept of scholarly quality. The general scholarly reaction to broad inquiries like the British Research Assessment Exercises has shown considerable disagreement both within and across disciplines as to the proper aims of knowledge in the social sciences and humanities. Indeed, the detailed reactions to the 2001 RAE say almost nothing at all about the substantive intellectual aims of the disciplines involved. Scientists and humanists alike view the evolution of "good knowledge" in exactly the same way: excellent research is whatever the discipline agrees is novel and important. In that data, it is uniformly believed - by humanists, natural scientists, and the vast majority of social scientists - that scholarly excellence cannot be measured by automated metrics, but can be recognized only by certified experts. Except for the statisticians, all respondents to the RAE commentary questionnaire argue that metrics are vitiated by bandwagon effects, rankings-induced inflation, arbitrary citation, and so on. (Abbott, forthcoming)

But if scholars agree that the only measure of excellent knowledge in the present is scholarly recognition, then - as I noted at the outset - we are without any clear substantive concept of the aims of knowledge. And without a clear substantive concept of the aims of knowledge, the optimality of digital tools cannot in fact be evaluated. If the only criteria of "improvement" are novelty and recognition, we are driven back to simple quantitative measures of scholarly output and citation, the very measures that scholars themselves have explicitly rejected again and again.

A different approach to this dilemma is to turn to the historical record. While we may be unable to agree on substantive criteria for knowledge in the present, we are certainly agreed that twentieth century scholarship was an extraordinary success. Thus, the kinds of quantitative measures that we find suspect in the present may be acceptable for an analysis of the past record. Even better, the theoretical issues were the same in the twentieth century as they are today; the debate was between "traditional research" and "new research technology." The project of universal knowledge databases searchable with high-quality indexes began around 1900. Throughout the twentieth century were developed a variety of universal reference tools, indexes, and so on. Yet surveys and ethnographies from at least the 1940s forward reveal that scholars actually made almost no use of the central research tools of the old library: the Union List of Serials with its guide to the exact periodical holdings of hundreds of scholarly libraries; the National Union Catalog with its details on book holdings; the various Wilson indexes to the scholarly periodical literature; and so on. Most scholars got most of their references either from specialized disciplinary reference works produced by colleagues or by hearsay (somebody told them about a good reference) or from the reference lists of other scholars' articles. (FN 1)

Thus, only one segment of the physical library's advanced tools was

actually used. That segment was *not* the universal indexes, but rather the particular, local views of literatures from particular scholarly vantage points: reference works produced by colleagues and specialist librarians. By contrast, the idea of a universal central library, unifying all the holdings of a given university and making them accessible to all users via a universal index, was the librarians' dream. Scholars vehemently opposed it. They wanted the books right next to their offices, in the departmental libraries where they had always been. In short, the Internet/Google project of dumping all "knowledge" into a single repository from which it can be extracted by "indexes" is about a century old. Scholars have always resisted it in practice - particularly the "universal index" part of it - and have devoted much time to developing their own very particular indexes to parts of the universal library.

In the event, the librarians won, of course, largely on cost grounds. All the major American universities built central libraries in the interwar period. But faculty insisted on getting offices within those libraries, because they still wanted to be right next to the books. Yet even in the central libraries, the survey evidence shows clearly that the majority of their trips to the stacks did not involve any detour through the card catalog. Faculty research practices were so contrary to librarians' image of what those practices ought to be that librarians routinely joked about something being "as rare as a faculty member who uses the subject index." Yet despite practices that seemed to librarians irrational and wrong-headed, scholars' vast productivity in the twentieth century testifies pretty clearly that they were right that centralization was probably unnecessary, whether or not they were also right that it was stupid.

In retrospect, one can see that the argument implicit in their practice made considerable theoretical sense. Each particular kind of user wants a particular kind of index, an index to the library from his or her disciplinary

or sub-disciplinary point of view. No universal index can serve such a user, because there are too many possible points of view to serve without losing focus in a welter of subject-headings. Indeed, on this conceptualization, every book in a library is a partial index to the library, and therefore it is sensible for scholars to get their references from the reference lists of other scholars, who have already created the indexes that, for all their faults, are of higher quality and focus than any universal or automated index can be. It's a question of empirical Bayesianism.

Yet late nineteenth and early twentieth century scholars *did* produce a version of the technology that dominates digital scholarship - the keyword index. Perhaps by looking at the impact of such old keyword indexes on past scholarship, we can get an idea of what we can expect from them today.

II - Concordances and Scholarship

I should begin with a clarification. True keywords are author-assigned tags. By limiting the number of these words, journals allow authors to send focal signals about their desired audiences. But in today's language, any word that appears in a text is called a keyword. An index that is an exhaustive list of all parts of a text using keywords is called - to give it its traditional name - a concordance, and thus what we today call keyword indexing should be called concordance indexing.

In the standard form, a concordance is an alphabetical list of all the words appearing in a work or a given body of work, together with an index to (and sometimes a listing of) all the units of the text (sentences, lines, verses) that contain those words. As in most digital systems, it was customary in concordances to leave out the simplest particles - articles, simple pronouns, and so on - but occasionally even these were included. For example, Alexander Crudden's 18th century concordance to the King James Bible includes

every word other than "a," "an," "the," and "and."

Concordances are quite old. By the High Middle Ages, there were lists locating all appearances of certain words in various sacred texts. As might be expected, the first complete concordances were to the Bible; concordances of the King James translation appeared shortly after its issue. Shakespeare's works were first concordanced in the eighteenth century. Manuals of quotations - a fundamental tool of oratory - were concordance-indexed by the middle nineteenth. (FN 2) Concordancing became a standard practice for literary figures in the late nineteenth century, and it is those concordances that provide a way to assess the impact of concordances on scholarship. Nearly all of these were concordances to poets and dramatists. I focus here on concordances to British poets.

It is not clear why some figures were concordanced and others not. A Chaucer concordance took fully 50 years and was published to great acclaim in 1927. By contrast, John Milton - who drew nearly as much scholarly attention as did Chaucer - was not concordanced until the computer era. One obvious issue was size of corpus, although Chaucer (1927), Shakespeare (18th century), Wordsworth (1911), Tennyson (1914), and Spenser (1915) all drew concordances, despite their voluminous work. Another possible reason for concordancing some authors rather than others seems to have been centrality to scholarship as it was then conceived. Early literary studies were mainly concerned with philological investigation, and thus emphasized early work like *Beowulf* (1911) and Chaucer (1927) (Graff 1987). But other works in the Old and Middle English periods languished - *Piers Plowman*, *Gawain and the Green Knight*, and so on. Nor is literary importance a clear cause; one is surprised to find Thomas Kyd (1906) concordanced before Christopher Marlowe (1911).

TABLE ONE

Concordances and Their Dates

Pope	1875
Cowper	1887
Burns	1889
Shelley	1895
Kyd	1906
Gray	1908
Beowulf	1911
Wordsworth	1911
Marlowe	1911ff
Tennyson	1914
Spenser	1915
Keats	1917
Browning	1925
Chaucer	1927
Herbert	1927
Herrick	1936
Collins	1939
Goldsmith	1940
Coleridge	1940
Housman	1940
Wyatt	1941
Donne	1941

Note: I have omitted two new Shakespeare concordances, one to the poems, appearing in 1874 and the other to Shakespeare's plays in 1894.

The list of concordances to major English literary figures between 1850 and 1945 is given in Table One. The dates are based on first editions of concordances as established in the *NUC Pre-1956 Imprints* and in Worldcat. As is evident, concordances were done in bursts, the first between 1906 and 1917, the second in the mid 1920s, the third in the late 1930s. Absent from this list are some very prominent figures, in particular Ben Jonson, William Blake, John Dryden, and Thomas Hardy, as well as John Milton, as already noted.

Concordances allow an easy way to test whether the existence of keyword indexing transforms scholarship. Scholars do not cite concordances, but we might expect that the availability of a concordance would tilt the work of

scholars toward a particular writer. Availability of a concordance would certainly facilitate work with a given writer, allowing one to see at a glance his use of images, his relative reliance on particular vocabularies, and so on. Indeed, one could argue that concordancing in American literary research was an attempt to industrialize the painstaking textual analysis that had dominated the German seminars where early American professionals had learned their literary studies. American scholars - Bliss Perry (1936:96ff.) is an example - complained bitterly of the mindless tedium of such work. Perhaps concordances were created to save time and allow scholars to turn to literary interpretation.

The archival evidence suggests that the people who did the concordances were indeed led by industrializers, but that those industrializers remained within the philological conception of literary scholarship characteristic of Germany. Albert S. Cook, the Yale professor who was original president of the Concordance Society, was a hardened philologist whose younger colleagues thought him an obstacle to the new literary studies (Canby 1947:213-216). Lane Cooper of Cornell was less conservative, but remained within the philological camp. So this first theory - that concordancing was done to take the Teutonic drudgery out of scholarship so that scholarship could move on to the more important matter of literary appreciation - seems unlikely on the archival evidence. It was after all the drudges who seem to have initiated the concordancing. However, it is not the origins but the results of concordancing that are most important for us. Even if the concordances were done by the drudges for the old purpose of philological research, it remains a plausible conjecture that concordances would be an aid to scholarship, and that the existence of a concordance would therefore encourage scholarship when it was accomplished. We should see more articles, books, and PhD dissertations on a poet after a concordance is done. I shall call this the facilitation

hypothesis. (FN 3)

There are two alternatives to the facilitation hypothesis. Perhaps by contrast concordances were produced by the same forces that produced bursts of scholarship. A glance at the dates of early PhDs on any major English poet reveals at once that there are fads in scholarship; the grouping of dissertations by year and indeed even by place is considerable. Particular seminars or particular research projects may have led to bursts of dissertations on one poet or another in one decade or at one particular university. It could well be that concordances were done as conveniences for the purposes of other research, and are therefore not so much facilitators of future scholarship as they are correlatives of ongoing scholarship.

But there might be another aspect to this correlative theory. It might be that the *doing* of a concordance produced the burst of scholarship rather than the *having* of a concordance. People doing concordances inevitably acquired great familiarity with the texts of an author. This would then facilitate their research on that author. Note that this theory suggests that we should expect a concordance to appear parallel to a burst of scholarly work. Indeed, since a concordance has to wait for the last co-worker's contribution (and then see considerable editing and production work), and hence is likely to follow the scholarship of most of its participants, it may be more likely that this mechanism would produce concordances immediately after a burst of scholarship. I shall call this the byproduct hypothesis.

A final hypothesis about the relation of concordances and scholarship is that the relation is arbitrary because the concordances were essentially amateur projects, having little to do with real scholarship. There is some suggestive historical evidence for this amateur hypothesis. Shakespeare's poetry was concordanced in 1874 by the wife of a Philadelphia lawyer, working alone and doing every single word (include "a" and "and". Not surprisingly, her son became a famous Shakespeare scholar.) Or again, Edwin Abbott,

schoolmaster and author of the brilliant *Flatland*, concordanced the works of Pope in 1875 because "every Englishman who wishes to use English words correctly will find help from knowing how Pope used them." (Abbott 1875:iii) John Bartlett, of *Bartlett's Quotations*, did a new Shakespeare concordance, ultimately published in 1896, which "has been prepared chiefly in leisure taken from active duties and from time to time has been delayed by other avocations." (Bartlett 1896:Note) Yale professor Albert Cook, mentioned earlier, did the concordance to Thomas Gray (1908) as a Christmas vacation amusement for himself and a few students and friends, quoting in his preface an anonymous writer in *Book News*: "No poet can be fully mastered without either dictionary or concordance. You have no grasp on a poet's use of words without one aid or the other." (Cook:1908:vii)

On the amateur hypothesis, we expect a relation to scholarship only if *having* a concordance really does help scholarship. If having a concordance is irrelevant to scholarship (the facilitation hypothesis is false) and doing a concordance doesn't have much impact either (the byproduct hypothesis is also false), then we won't see much scholarship associated with concordances one way or the other.

In short, there are three general theories about the relation of concordances to scholarship, with clear and differing implications for the timing patterns of scholarly production. If having a concordance facilitates scholarship, then we expect there to be more scholarship after a concordance than before. If doing a concordance facilitates scholarship or if concordances proceed as a part of ongoing research, then we expect bursts of scholarship either contemporaneous with the publication of a concordance or, given the different time structures, before the publication of a concordance. If amateurs do concordances and having them is useful, then we will see the first pattern. But if scholars don't do concordances (thereby acquiring useful

familiarity with their subjects) *and* concordances don't facilitate scholarship, then rates of scholarship will not be particularly associated with concordances at all.

III Concordances and Scholarship - Empirical Analysis

Between 1850 and the onset of computerized concordancing immediately after the Second World War, twenty-two concordances were published to major figures in English literature, as noted in Table One. Fortunately, we possess a complete listing of all PhD degrees in English literature from 1865 to 1964 (McNamee 1968). This covers all universities in Germany, Great Britain, and the U. S. Moreover, the work's preface makes clear that this listing is as close to definitive as is humanly possible, having been based on hand-checking of records by most of the universities involved. It is certainly far more accurate than such general compilations as those of the American Council on Education and the national data used for the *Historical Statistics of the United States*. (Because of the historical period, such later lists as *Doctoral Dissertations in English* are useless, and, indeed, McNamee's checking with universities proved that many such lists were quite inaccurate.) Luckily for our purposes, this listing is also subject-indexed. There are short-comings, to be sure; dissertations are listed under only one heading, and dissertations on multiple figures go automatically under the first one. But as historical data go, McNamee's listing is very good.

There is a similar work, but of much lower quality, for BA and MA theses in English literature (Howard 1973). This work makes no pretense to completeness. It was developed by concatenating several prior indexes and augmenting these with personal research and extensive correspondence. (Unlike McNamee, who had Office of Education funding, Dr. Howard produced her work as a labor of love.) However, there is little reason to think that its sampling procedure produces biases that are differential across our three hypotheses,

as long as we are careful to avoid making detailed comparisons across authors (some of the originating bibliographies concerned particular authors). The list of cooperating universities and colleges makes it clear that the sample of the work is very wide, and that there is little reason to expect systematic biases that will undercut general measures.

On articles and books themselves, the data are more complicated. From 1921 on, we have the *Bibliography of English Language and Literature* (BELL), which remained in a single consistent format throughout the period here considered. It is true that the issues for the war years (1939-1945) were issued only in the 1950s, which could imply that the analysis was more complete because work had had time to enter various bibliographies, or that it was less complete because it had to be done in parallel with current issues which were, in most scholars' eyes, more important). But BELL seems to have been collected by national sub-editors responsible for surveying the contents pages of a particular set of journals, so retrospective recreation does not seem to pose a problem.

On the other hand, there is no such source for the period before 1921, although this includes half the concordances of interest. A number of expedients are possible. The general periodical indexes (*Poole's*, *Nineteenth Century Reader's Guide*, *Reader's Guide to Periodical Literature*, and the *International Index*) cover only a small portion of the scholarly literature. They do not include books (as does BELL). *Poole's* does not produce consistent dates, and its computer searchable versions are not effectively subject-indexed at the level needed (there are about seventy different headings for Keats, for example.) While the Wilson indexes (the other three listed above) do have effective subject indexing, a little work with them shows that they do not cover enough material to provide any real evidence of rates of scholarly work. There are too few articles to estimate rates.

The only recourse before 1921 is therefore JSTOR. Unfortunately, JSTOR has no subject indexing whatever. It is only keyword - that is, concordance - indexed. An article entitled "The Ode and its Urn" will not be identified with Keats in the JSTOR indexes except by the appearance of Keats's name on many pages in the article, which necessitates searching on the appearance of "Keats" in the article, checking the number of times it appears, and then establishing whether the article as a whole is about Keats's work. This would be an immense and error-ridden task. Luckily, there is little reason to think that scholarship influenced by concordances is more or less likely to put the name of the poet involved in its title than is scholarship not influenced by concordances. That is, again there is no *differential* bias of these tools across our three mechanisms. We can thus use the JSTOR indexes searching titles on the subject poet's name. In cases where the name is a common one (Collins or Gray, for example) we have to add restrictions: first requiring that the articles also include the word "poet" in the text and second requiring the relevant first name. (Otherwise "Pope" produces articles about the Vatican as well as about the eighteenth century poet laureate.) Recall that all comparisons will be within poet, and all scholarship is retrieved with one command string, so any biases arise only through different behavior by concordance-influenced and concordance-non-influenced scholarship with respect to that search strategy. Such bias seems very implausible.

A. Doctoral Dissertations

Doctoral dissertations are a high standard to set, since there are relatively few of them. As of 1920 there were about 25 American dissertations in all of letters (not just English literature, but all literatures) in a given year. By 1930, the figure was about 90, by 1940 about 175, and by 1950 240. Adding in the German dissertations helps - and there is no reason that young German scholars writing on English poets would not have used English-

language concordances. But even so, there were, for example, no dissertations before 1900 on Donne and Coleridge, and only one or two on figures like Wordsworth, Shelley, and Tennyson. (By contrast, there were quite a few on *Beowulf*, on Chaucer, and on Spenser; early literary studies, as noted earlier, were largely philological.) Thus it is only with the concordances after 1910 that we can tell anything about the rates of dissertation writing. Table Two presents, for the concordances from *Beowulf* (1911) onward, the total number of dissertations (for the US and overall), the number written before five years pre-concordance, the number written in the five years leading up to the concordance, and the number in the concordance year and the subsequent four years.

I have chosen the five-year interval because it is wide enough to capture a reasonable number of events. Note however that the expansion of PhD production - particularly marked in the 1920s and the period after 1945 - means that the raw figures (with the exception of one set noted below) are strongly biased towards accepting the facilitation hypothesis; there should be more scholarship after a concordance simply because scholarship itself was rapidly expanding. There were 1.75 times as many literature PhDs in the late 1920s as in the early 1920s, 1.75 times as many in the early 1930s as in the late 1920s, and 1.32 times as many in the late 1930s as in the early 1930s. Even in the early 1940s, there were slightly more PhDs (5% more) than in the late 1930s, despite a rapid fall after 1943. Thus, throughout this period, even if the concordances had no effect on scholarship we should expect the number of PhDs in the concordance year and its four successors to exceed those in the preceding five years by a factor of at least 1.7 up through 1935, falling to 1.3 and then eventually 1.05 in the late 1930s and early 1940s.

TABLE TWO

Concordances and PhD Production

Concordance	US PhDs only				US, Britain, Germany PhDs			
	N	TO C-5	C-5/C-1	C/C+4	TO C-5	C-5/C-1	C/C+4	
Beowulf 1911	21	1	0	0	1	0	0	
Wordsworth 1911	80	2	2	1	5	5	2	
Marlowe 1911ff	24	1	0	0	7	7	2	
Tennyson 1914	31	0	0	1	6	4	4	
Spenser 1915	104	3	3	3	15	6	4	
Keats 1917	42	0	0	0	4	0	0	
Browning 1925	60	3	2	2	10	7	5	
Chaucer 1927	170	35	7	7	74	9	10	
Herbert 1927	8	0	0	0	0	0	0	
Herrick 1936	6	0	0	1	1	1	1	
Collins 1939	1	0	0	0	0	0	0	
Goldsmith 1940	19	6	2	2	14	2	2	
Coleridge 1940	43	5	3	3	17	6	3	
Housman 1940	2	0	0	0	0	0	0	
Wyatt 1941	9	1	0	1	3	0	1	
Donne 1941	49	10	7	2	14	7	4	
Column Sums		67	26	25	173	56	42	

There is in this table no evidence for the facilitation hypothesis. The quinquennia immediately before a concordance produce as many dissertations as do those immediately after. Indeed, if there is any direction here, it is the reverse of that expected. PhD production in the post concordance period is roughly the same at that of the immediate pre-concordance period, even though the rapid expansion of academia leads us to expect a quite substantial increase even in the absence of facilitation. The evidence against the facilitation hypothesis at the PhD level is thus strong. (FN 4)

The right hand columns report the same figures but include the British and German PhDs. Here the evidence is even stronger against the facilitation hypothesis, although we should remember that the European systems were not rapidly expanding during this period as were the American ones. Yet still, this evidence seems strongly against the facilitation argument.

Not shown in the table, but extremely evident in the sources, is the occasional grouping of dissertations in narrow bursts, both within time periods and occasionally within universities. Harvard produced 20 of the 35 American Chaucer PhDs before 1920. Ten of those came in eight years: 1905, 2 in 1906, 1907, 1908, 1909, 4 in 1911, 1912, 1913. Yale produced six Chaucer PhDs between 1934 and 1939, after having produced only three in its entire history before that time. Chicago produced 5 Chaucer PhDs in the 1930s, hardly surprising since faculty members John Manly and Edith Rickert were producing the definitive Chaucer manuscript at the time. As for Spenser, nearly half (42%) of Spenser PhDs were produced within two years of another Spenser PhD at the same institution.

Such bursts could have been driven by newly available library or manuscript materials at particularly universities. But since these bursts of PhDs are not consistently over time at any one institution (although Yale and Harvard are quite preeminent), it is probably not the case that this mechanism had much effect. Rather, seminars, research projects, and particular faculty probably drove the phenomenon of dissertation-bunching.

The raw PhD information and archival data provide some substantive validation for a general version of the byproduct hypothesis, but, as we shall see, particular bursts of PhDs at particular universities do not seem important. If we look at the raw data for the more studied poets, looking for possible bases for "byproduct" production of concordances, we find that such bases probably exist for:

Wordsworth - dissertations in 1906, 1907, 1908, 1909, and 1910 leading up to the concordance (and another dissertation) in 1911.

Tennyson - dissertations in 1909, 1912, and 1913 (2) leading up to the concordance (and another two dissertations) in 1914.

Spenser - dissertations in 1910, 1911, 1912 (3), and 1913 leading up to the concordance in 1915.

Browning - dissertations in 1921 (2), 1922, 1923, and 1924 (3) leading

up to the concordance in 1925.

Chaucer - dissertations in 1923 (2), 1924 (3), 1925, and 1926 (2) leading up to the concordance (and another dissertation) in 1927.

Coleridge - dissertations in 1935 (2), 1937, 1938, and 1939 (2) leading up to the concordance in 1940.

Donne - dissertations in 1936, 1937, 1939 (4), and 1940 leading up to the concordance (and another two dissertations) in 1941

But in only one case do we see in the immediate pre-concordance period a concentration at a single university (on Chaucer at Stanford - 3 dissertations in 1924 and one in 1926). Moreover, in nearly all cases at least half of the dissertations involved are German. Thus the bursts are across the field, rather than in one place. Yet in most cases the animator of the concordance was a scholar: Lane Cooper (Cornell) on Wordsworth, Charles Osgood (Princeton) on Spenser, Leslie Broughton (Cornell) on Browning and (in part) Keats, John Tatlock (Stanford 1915-25, Harvard 1925-1929) on Chaucer, and Homer Combs (Northwestern) on Donne. Amateurs seem to have led only the Tennyson and Coleridge concordances.

The PhD information, therefore, pretty strong pushes us towards the byproduct hypothesis. Not only do the numbers clearly favor it, the substantive information we have indicates that the concordances probably were often produced as conveniences to ongoing academic work and as disciplines for students, the creation of a concordance itself being a kind of formal training via immersion in the text.

B. Concordances and Articles in JSTOR

We now turn to articles in JSTOR. These have been sampled as noted above. Table Three presents the results. The article figures are for the same three periods: from the beginning up to five years pre-concordance, from five years pre-concordance through the last pre-concordance year, then from the concordance year to the fourth subsequent year. Here the column sums reveal a slight

increase from before the concordance to after it. But the increase is less than the increase in simple PhD numbers: the simple numerical increase in PhDs over any successive five year periods was at all times (except the early 1940s) more than the 16% increase seen here. So this evidence too probably weighs against the facilitation hypothesis, although not as strongly as does Table Two.

TABLE THREE

Concordances and Articles in JSTOR

	Concordance	N	TO C-5	C-5/C-1	C/C+4
Pope	1875	110	0	1	0
Cowper	1887	20	0	0	0
Burns	1889	39	0	1	0
Shelley	1895	195	1	4	2
Kyd	1906	6	0	2	0
Gray	1908	29	0	1	0
Beowulf	1911	188	15	7	9
Wordsworth	1911	228	6	6	4
Marlowe	1911ff	71	0	10	3
Tennyson	1914	104	29	5	3
Spenser	1915	283	26	20	9
Keats	1917	160	12	4	5
Browning	1925	134	36	10	10
Chaucer	1927	682	205	40	66
Herbert	1927	15	3	1	0
Herrick	1936	24	10	0	1
Collins	1939	11	7	0	2
Goldsmith	1940	53	33	9	5
Coleridge	1940	132	54	11	10
Housman	1940	31	1	2	9
Wyatt	1941	25	9	1	1
Donne	1941	119	36	9	29
Column Sums			483	144	168

The major increase is confined to three authors, Chaucer, Donne, and Housman. But inspection of article titles reveals that the Housman increase reflects the poet's death in 1936, which spawned a wave of Housman studies under the leadership of his brother Laurence. One surmises that the concordance itself was stimulated by the poet's death, but its introduction is

silent on that subject (Hyder 1940). As for Chaucer, the increase in JSTOR articles probably originated not in the concordance of 1927 but in the University of Chicago's monumental Chaucer project (1924-1940), which would produce the definitive reconciliation of all existing Chaucer texts (Manly and Rickert 1940). There is no such obvious explanation for the increase in work on Donne. T. S. Eliot's celebrated essay on the metaphysical poets dates from 1921, and there was to my knowledge no major Donne project around 1940. But in the other two cases there are obvious special circumstances.

In summary, while this table does not contradict the facilitation argument as strongly as does Table Two, it nonetheless favors the byproduct argument slightly more than it does the facilitation argument. Both tables, moreover, underscore that the effects are not consistent across authors and that particular events - seminars, colloquia, research projects, disciplinary waves of enthusiasm - almost certainly overwhelm any possible concordance effects.

The JSTOR data, because of their breadth in time, allow us to consider - albeit indirectly - a possible alternative mechanism explaining the relation between concordances and output. This is the possible differential distribution of the concordances themselves. Some were undoubtedly more available than others. There is, of course, no general data on past library holdings at any given point in time. However, we can at least consider the relation between current holdings of concordances and totals of JSTOR articles. If there is a strong relation between the two, distribution might matter, although of course the JSTOR data is dominated by the recent period because of PhD expansion. Table Four presents for each poet the date of the concordance, the number of print copies currently recorded in Worldcat (in a few cases, separate volumes are counted as separate items), the number of JSTOR articles, and the ratio of JSTOR articles to Worldcat holdings. The table is ordered by the last of these. It is at once clear that while there is

some correlation, it is pretty mild and probably reflects exogenous factors rather than a causal relation between concordance availability and output. Nonetheless, it is interesting that the top three authors, in terms of this ratio, are Spenser, *Beowulf*, and Chaucer, the philologists' central subjects. Again, we see that concordances may have served that particular form of analysis.

TABLE FOUR

Concordance Distribution and JSTOR Production

Poet	Concordance	Copies	JSTOR	Ratio
Kyd	1906	126	6	0.05
Collins	1939	223	11	0.05
Wyatt	1941	344	25	0.07
Burns	1889	399	39	0.10
Herbert	1927	150	15	0.10
Herrick	1936	204	24	0.12
Gray	1908	200	29	0.15
Cowper	1887	104	20	0.19
Donne	1941	589	119	0.20
Tennyson	1914	386	104	0.27
Housman	1940	104	31	0.30
Pope	1875	305	110	0.36
Coleridge	1940	347	132	0.38
Keats	1917	363	160	0.44
Browning	1925	289	134	0.46
Shelley	1895	384	195	0.51
Goldsmith	1940	104	53	0.51
Marlowe	1911ff	130	71	0.55
Wordsworth	1911	402	228	0.57
Spenser	1915	358	283	0.79
Beowulf	1911	172	188	1.09
Chaucer	1927	404	682	1.69

More generally, however, the "availability" hypothesis is rejected by the distribution data alone. The lowest number of concordances copies available in 2010 was 104, for Housman, Cowper, and Goldsmith. But throughout the first half of the twentieth century, PhD level scholarship was intensely centralized in the United States; in many disciplines, a third of all PhDs were granted by the top five departments as late as the 1950s. (Abbott, in press) It can thus

be safely assumed that during the period here studied every PhD level scholar in the US did his dissertation at a university whose library possessed all of these concordances.

C. General Publications after 1920

The list of publications contained in BELL provides another useful body of evidence. The source exists only from 1921, however, so only later concordances can be studied. On the positive side, BELL includes books as well as articles. This however is a double-edged sword, since BELL includes popular books as well as scholarly ones. With no clear ground for distinguishing the two and with the further knowledge that any such ground would change over time, we must simply accept the editors' own judgment, as evidenced in their including works in this compilation, remembering that there is some popular work in here as well as scholarship. (Moreover, the amount of popular work varies over time: the editors' introductions often speak of trying to cut down on less professional material to make room for the burgeoning research literature.) BELL also includes short comments raising or answering particular questions as well as new editions of standard works. (However, all reviews of any new edition are counted as one item with that revision, as they are in BELL itself.) It is thus a worrisome source, although the best available.

TABLE FIVE

Concordances and articles in *Bibliography of English Language and Literature*

Relative to concordance year

		RAW		CORRECTED	
		+++++		+++++	
		c - 5	c + 5	c - 5	c + 5
Browning	1925	62	73	31.2	24.5
Chaucer	1927	150	232	55.6	65.3
Herbert	1927	2	4	.7	1.3
Herrick	1936	13	7	3.0	1.5
Collins	1939	9	11	1.9	2.6
Goldsmith	1940	28	21	5.9	6.8
Coleridge	1940	100	83	20.9	25.0
Housman	1940	52	47	11.1	18.3
Wyatt	1941	3	2	.6	.6
Donne	1941	53	65	10.9	22.9
Column sums		472	545	141.8	169.0
			+15%		+19%

Table Five shows the BELL data for the authors whose work was concordanced after 1920. The first two columns give a pair of raw counts: the first for the five years immediately prior to the concordance, the second for the concordance year and the four subsequent years. These are numbers of separate entries in BELL. As in Table Three there is an increase of about 15% from pre-concordance to post-concordance, most of it driven by the 50% increase in works on Chaucer. Given the expansion of PhD production and the enormous output of the Chaucer project, this evidence too provides little support for a facilitation argument, and indeed seems to favor the byproduct argument.

The second set of figures in the first panel *controls* for the total number of items appearing in BELL in a given year (that is, it controls for the total number of entries across all literary subjects). It thus provides figures uninfluenced by the expansion and shrinkage of literary academia. The figure shown is a function of the yearly portion that items by the given poet

contribute to the total of all items. For example, in 1922, there were 28 Chaucer items and 2443 items total. Taking these figures for all five years 1922 to 1926, we have 28/2443, 38/3088, 23/2284, 23/2041, and 38/2727. The sum of these fractions (0.0095, 0.0125, 0.0101, 0.0096, and 0.0139) is .0556. I have multiplied this by 1000 to give a more readable figure. (The figure presented is thus 50 times the actual annual average: Chaucer items were 55.6 / 50 = 1.11 percent of all literary items in BELL in this period.)

This correction provides the first positive evidence for the facilitation hypothesis. The quotient of the before and after column sums indicates a rise of 19%, which originates in substantial increases for Chaucer and Coleridge and massive increases for Donne and Housman, all of which set off the striking decrease of Browning. The Chaucer increase, as noted, probably reflects the increasing presence of a massive project. Housman, as earlier noted, was the subject of a small-scale industry from his death in 1936 onward. The actual number of Housman articles fell after the publication of the concordance, but the total coverage of BELL fell even more. Indeed, it is important to note that all volumes of BELL from 1939 to 1945 were done retrospectively after the war, so there is a more purposive kind of sampling involved in the creation of the underlying data itself. This could be influencing the figures somewhat, although there is little reason to think that the influence would be differential between authors, which is the only influence likely to bias these results.

As for the other changes here, there is no obvious account of the sudden expansion in writing about Donne nor the considerable increase in work on Coleridge. Nor is there an obvious explanation of the decline in Browning studies. The important fact is thus that the BELL data do provide some evidence for the facilitation theory.

TABLE SIX

Concordances and articles in Bibliography of English Language and Literature

Absolute Peaks

		RAW	CORRECTED
		+++++	+++++
Browning	1925	c + 13	c - 5
Chaucer	1927	c + 10	c - 5
Herbert	1927	c + 7	c + 16, c+17
Herrick	1936	c - 7, c-3, c-1	c - 7
Collins	1939	c - 9	c - 9
Goldsmith	1940	c - 12, c-11	c - 11
Coleridge	1940	c - 6	c - 6
Housman	1940	c + 4	c + 3, c+4
Wyatt	1941	c - 7	c - 20
Donne	1941	c - 10	c + 2

Table Six reports the date of the single highest year's output for any given author in the BELL data, both in the raw statistics and in the period-corrected ones. The raw statistics are included for intra-author comparison only, since we are comparing output over a twenty-year period that saw large increases in total output, so single year figures mean little. The corrected figures show that in most cases the highest single year *precedes* the concordance, the exceptions being Housman and the metaphysical poets Donne and Herbert, both of whom peak in the early 1940s. Perhaps there was a professional focus on the metaphysical poets during the war years.

In summary then, the evidence from the BELL data is somewhat positive for the facilitation hypothesis. Although the raw figures continue the negative verdict (given the expansion of academia), there does seem a rise in period-controlled output immediately after a concordance, although two of those figures seem obviously attributable to alternative factors. The peak output figures give pause for the facilitation claims, for they show that the major peaks typically precede the dates of concordances. Moreover, the facilitation effect is not consistent across authors, but rather depends mostly on four

writers of the ten. At the same time, of the top five authors, all but one shows a post-concordance increase on the period-controlled figure. Overall, the data incline to facilitation.

D. BA and MA Theses.

My final data concern BA and MA theses and are presented in Table Seven. Although these show some important changes from the PhD data, their overall message is the same. Again there is a problem with small numbers in the early years. Although there are of course more MA papers than dissertations at any given time, until the Wordsworth concordance of 1911 there are in this data no BAs or MAs on a concordanced figure within five years (in either direction) of the relevant concordance. Thus the analysis does not include earlier concordances, although there are, of course, scattered BAs and MAs going back to 1894, when this data starts.

TABLE SEVEN

Concordances and BA and MA Theses					
+++++					
Concordance	N	TO C-5	C-5/C-1	C/C+4	
Wordsworth 1911	124	6	5	3	
Marlowe 1911ff	92	0	0	0	
Tennyson 1914	145	11	2	4	
Spenser 1915	171	9	9	5	
Keats 1917	133	4	0	0	
Browning 1925	232	18	5	3	
Chaucer 1927	214	13	3	12	
Herbert 1927	56	0	0	2	
Herrick 1936	28	2	4	6	
Collins 1939	6	0	2	1	
Goldsmith 1940	50	6	9	4	
Coleridge 1940	114	13	6	9	
Housman 1940	23	2	0	1	
Wyatt 1941					
Donne 1941	169	8	16	9	
Column Sums		92	59	69	

Glancing at the total figures, we see that Tennyson and especially

Browning with their heavily philosophical content are more important subjects at this lower level, and the prominence of Donne and even Herbert bespeaks the popularity of the metaphysicals in the 1950s and 1960s. It is also important to recall again the rapid expansion of academia in the period after 1910. Although there are no national data on BAs and MAs comparable to the PhD numbers (historical BA and MA data are not field specific), we can sum these lists by year across all these concordanced poets. By this measure, we find that there were about 1.5 times as many theses in the early 1920s as in the late 1910s, about twice as many in the late 1920s as in the early 1920s, about 1.5 times as many in the early 1930s as in the late 1920s, and about 1.25 times as many in the late 1930s as in the early 1930s. There were, however, only .7 times as many in the early 1940s as in the late 1930s.

These figures mean that in Table Seven the rapid growth in Chaucer theses is probably due to simple expansion of the college and university student body, since the Chaucer concordance came during the heyday of that expansion. (It is of course very unlikely that these theses arose in the Chicago Chaucer Project.) By contrast, the figures for Donne and Coleridge would undoubtedly have been higher if academia had not contracted in the 1940s. These effects roughly cancel each other, leaving us with a very slight expansion in theses in the postconcordance periods, on average. This is probably due to simple expansion - even around 1910 BA and MA level academia was growing steadily. There is in any case not an expansion strong enough to suggest a decisive facilitation effect, although it seems more positive than the PhD evidence in Table Two. One might hypothesize that BA and MA level students would have more use for such concordances than would advanced scholars who knew the underlying texts far better. Just as quotation dictionaries are mainly useful for non-experts giving occasional orations, so too might concordances be mainly useful for non-experts studying poets in the course of liberal education.

IV Conclusion

We had three hypotheses about the impact of concordances. Under the facilitation hypothesis, concordances arose fortuitously and then by their very presence shaped the research process, by providing a tool that would attract scholarly attention and facilitate scholarly work. Under the byproduct hypothesis, concordances arose as another expression of general bursts of scholarly interest that would also bear fruit in PhDs, articles, and BA and MA theses. Under the amateur hypothesis, concordances were largely irrelevant to scholarship, generated by different processes and serving different interests - private enjoyment, oratorical munitions, and so on. Empirically, the first of these hypotheses would be shown by substantial bursts of scholarly activity after the dates of concordances, net of system expansion. The second would be shown by bursts of scholarly activity concurrent with or prior to a concordance. The third would not be shown by any particular timing, but rather by ambiguous timing evidence on the first two hypotheses coupled with documentary evidence to showing the origins of concordances in such alternative interests.

The evidence seems first of all to reject the facilitation hypothesis. It is decisively rejected by the PhD data of Table Two and also rejected - although less strongly - by the JSTOR data of Table Three and the peak years BELL data of Table Six. The lone quantitative support comes from the BELL data of Table Five, on a truncated sample, with alternative accounts for at least two of the poets whose figures are responsible for the overall change. The BA and MA data of Table Seven suggests the possibility of a facilitation effect, but only below the highest level of scholarship. In sum, it is hard to maintain facilitation as a major scholarly effect, although we might think concordances facilitate work in more popular settings as well as at lower academic levels.

Conversely, the data seems to favor the byproduct hypothesis, particularly once we begin to turn to archival and historical information to interpret the quantitative results. Table Two provides clear evidence for the byproduct theory, and the detailed investigation of concordance leaders and bursts of scholarship seems to suggest a substantive basis for this theory. Table Three also supports the byproduct hypothesis, although less strongly. Both Tables Two and Three underscore the erratic nature of PhD and article production. By contrast, Table Five - particularly in the output controlled figures - seems to counter the byproduct hypothesis. This could however reflect the fact that dissertation production is a stricter standard than is publication, since BELL includes not only formal scholarship, but also new popular editions, general market monographs, popular treatments, and so on. Table Six suggests that the highest bursts of production tend to antedate the concordances, which favors the byproduct hypothesis. Finally, Table Seven's data on BA and MA theses, like Table Three's on JSTOR articles, favors the byproduct hypothesis, but more likely points to a lower level facilitation effect.

A conservative conclusion from this data would be that there is no strong and consistent evidence that concordances tended to produce or facilitate bursts of scholarship on the authors concordanced. The evidence seems to strongly suggest, but not confirm, that the causality runs the other way; concordances are one of the many by-products of bursts of scholarship already under way. It is striking, too, that the evidence against facilitation is strongest for the most expert level of production - PhD dissertations. The article and BA/MA data are more equivocal, and the BELL data, with its admixture of popular material, is the only data to favor the facilitation hypothesis. Facilitation, if it exists, happens below elite scholarship.

The amateur hypothesis cannot be ruled out without more detailed investigation of the archival data. It is clear that concordancing involved

many amateur workers. That Cornell faculty wives were set to work by Lane Cooper and his colleagues was a common joke in Ithaca. Thus, the first indications from the archival data are that concordancing was something organized by scholars but often done by amateurs. This kind of distributed scholarship was common in literary studies: the Oxford English Dictionary had after all been produced by such means. Indeed, the period before 1930 seems to have seen strong collaboration between experts and amateurs in many fields, from astronomy to sociology and literary studies.

For the present, it seems that we can conclude that concordances probably arise as one aspect of a general increase of interest in a poet and are produced under the leadership of scholars who see them as useful both as tools and as training, and who involve amateurs actively in the process. Concordances probably do not fundamentally shape scholarship, nor do they seem to facilitate it on a really noticeable level. At the same time, they may have considerable facilitative implications for less expert audiences: for first-level students and for pure amateurs.

These conclusions have distinct implications for our understanding of the modern research world, which is dominated by the keyword indexed tools of the internet. Such tools make certain aspects of scholarship easier than before. In a few cases, they make things possible that were not possible before. But the present results suggest that they probably will not revolutionize real scholarship, any more than the Chaucer concordance revolutionized the search for a consensually-agreed-upon Chaucer manuscript. Quite the contrary, the former presupposed the latter. (The reader will have noticed that the Chaucer concordance was perforce done using a text that was not definitive and that was replaced shortly after the concordance's completion.)

Finally, a more general word about the implications of these results for research in the humanities and the humanistic social sciences (HHSS). As I have argued elsewhere (Abbott 2008), it is quite evident that the knowledge

mode of HHSS is different from that in the sciences and that the general knowledge project of HHSS is quite unlike that of the sciences. The sciences believe in a search for truth and in the ability of nature, rightly interpreted, to guide them on their search for truth by rejecting incorrect theories. In HHSS, we are pursuing interpretations rather than truth, and no one feels that there is an ultimate or final interpretation of any novel, sonata, or painting. Improvement of research probably has more to do with the plenitude or density of interpretations than with the discovery of some "truth," and it is quite unclear, given that project, whether rapidity of search affects the outcome sought - improvement of plenitude - in any particular way. Indeed, it was the move away from philology's highly scientific approach that produced the efflorescence of English literary scholarship in the twentieth century.

It seems more likely that the impact of keyword indexing will be felt further down the intellectual ladder, at the level of what we might call "amateur experts." There are hints in the analysis above that these groups - undergraduates and MA writers as well as pure amateurs - have actually been the main beneficiaries of keyword indexes. It could be that the main utility of keyword indexes is to provide new possibilities for those amateurs. The move of scholarship to a largely professional basis in the 1920s and 1930s drove the amateurs out of most scholarly societies, as I have shown elsewhere (Abbott in press.) Perhaps keyword indexes will provide them with a way back in.

As for the scholars themselves, there is little indication in these data that keyword indexes will revolutionize their inquiries. There will be a return to philology, since keyword indexes permit tracing of the history of words on a level never before possible. But the literary scholarship of the twentieth century, if it accomplished nothing else, showed beyond the shadow

of a doubt that the histories of words and the histories of meanings are not the same thing. Understanding the relation between the two requires an individual mastery of a particular time, place, and language that is not going to be automated any time soon.

FOOTNOTES

FN 1. The following paragraphs are based on the study reported in Abbott (in press.) Detailed sources can be found there.

FN 2. Quotation manuals are themselves very old tool, deriving from medieval manuals of learning. See Wright 1931.

FN 3. I have not yet completed the archival side of this project. Statements about the history or concordancing are based on preliminary reading of relevant material from the Lane Cooper Papers at Cornell and Rutgers, and the Albert Cook papers at Yale. See also Abbott (in press.)

FN 4. The PhD figures are from the Millenium Edition of the Historical Statistics of the United States, 2:450, series Bc610. The reader may be surprised that I am presenting and analyzing data as simple counts. But the erratic nature of the data make it quite clear that more formal methods - hazard-rate modeling, for example - are ruled out because of the assumptions necessary. Note also that no uniform expansion rate (of PhDs) can be applied as an "expectation," since the concordances happen at many different dates. One could use the expansion rates noted above to estimate for each poet the number of post-concordance dissertations expected on mere grounds of expansion, but this would create an air of false exactitude. The same goes for poisson and negative binomial regression, which could be applied here, given that the dates of all dissertations are known. The data are too sparse for such methods to make much sense. Graduate student Kinga Makovi at Columbia raised the interesting question of whether the apparent relative decline of PhDs after a concordance might not be caused by increasing subdivision of dissertation topics, and in particular by a secular turn to dissertations on other poets. This mechanism might obtain much later in the period: by the

1960s dissertation topics were indeed being exhausted at a heroic rate. But at this time, there were still many open topics even for such well-studied figures as Chaucer. As the table shows, there had been 170 Chaucer dissertations as of 1964, and only 49 of these came before 1931, four years post-concordance. It seems unlikely that subdivision is driving the results.

CONCORDANCES (By Date)

- Pope Abbott, E. 1875. *A Concordance to the Works of Alexander Pope*
New York: Appleton
- Cowper Neve, John. 1887. *A Concordance to the Poetical Works of
William Cowper*. London: Sampson, Lowmarston, Searle, and
Rivington.
- Burns Reid, J. B. 1889. *A Complete Word and Phrase Concordance to the
Poems and Songs of Robert Burns*. Glasgow: Kerr and
Richardson.
- Shelley Ellis, F. S. 1892. *A Lexical Concordance to the Works of Percy
Bysshe Shelley*. London: Bernard Quaritsch.
- Kyd Crawford, C. 1906. *A Concordance to the Works of Thomas Kyd*.
Louvain: Uystpruyst.
- Gray Cook, A. S. 1908. *A Concordance to the English Poems of Thomas
Gray*. Boston: Houghton-Mifflin.
- Beowulf Cook, A. S. 1911. *A Concordance to Beowulf*.
Ithaca: Cornell University Press.
- Wordsworth Cooper, L. 1911. *A Concordance to the Poems of William
Wordsworth*. London: Smith Elder.
- Marlowe Crawford, C. 1911ff. *The Marlowe Concordance*.
Louvain: Uystpruyst.
- Tennyson Baker, A. E. 1914. *A Concordance to the Poetical and Dramatic
Works of Alfred, Lord Tennyson*.
London: K. Paul, Trench, Trubner.
- Spenser Osgood, C. 1915. *A Concordance to the Poems of Edmund Spenser*.
Washington: Carnegie Institution.
- Keats Baldwin, D. L. et al. 1917. *A Concordance to the Poems of John
Keats*. Washington: Carnegie Institution.
- Browning Broughton, L. N., and B. F. Stelter. 1925. *A Concordance to the
Poems of Robert Browning*. New York: G. E. Stechert.
- Chaucer Tatlock, J. S. P., and A. G. Kennedy. 1927. *A Concordance to
the Complete Works of Geoffrey Chaucer and to the Romaunt
of the Rose*. Washington: Carnegie Institution.
- Herbert Mann, C. 1927. *A Concordance to the English Poems of George
Herbert*. Boston: Houghton-Mifflin.
- Herrick McLeod, M. 1936. *A Concordance to the Poems of Robert Herrick*.

New York: Oxford.

- Collins Booth, B. A., and C. E. Jones. 1939. *A Concordance to the Poetical Works of William Collins*. Berkeley: University of California.
- Goldsmith Paden W. D. and C. K. Hyder. 1940. *A Concordance to the Poems of Oliver Goldsmith*. Lawrence KA: University of Kansas.
- Coleridge Logan, Sister E. 1940. *A Concordance to the Poems of Samuel Taylor Coleridge*. St. Mary in the Woods: Privately Printed.
- Housman Hyder, C. K. 1940. *A Concordance to the poems of A. E. Housman*. Lawrence KA: University of Kansas.
- Wyatt Hangen, E. C. 1941. *A Concordance to the Complete Poetical Works of Sir Thomas Wyatt*. Chicago: University of Chicago Press.
- Donne Combs, H. C., and Z. R. Sullens. 1941. *A Concordance to the English Poems of John Donne*. Chicago: Packard.

Note: I have omitted two new Shakespeare concordances, one to the poems, appearing in 1874 and the other to Shakespeare's plays in 1894.

REFERENCES

- Abbott, A. 2008. "The Traditional Future: A Computational Theory of Library Research." *College and Research Libraries* 69:524-545.
- . in press. "Library Research Infrastructure for Humanistic and Social Scientific Scholarship in America in the Twentieth Century." In M. Lamont, C. Camic, and N. Gross, eds. *Social Knowledge in the Making* Chicago: University of Chicago Press.
- . forthcoming. "The Excellence of IT" In M. Herbst and L. Greenfeld, eds. *The Institution of Science and the Science of Institutions*. New Brunswick NJ: Transaction.
- Abbott, E. A. 1875. *A Concordance to the Works of Alexander Pope*. New York: Appleton
- Bartlett, J. 1896. *A New and Complete Concordance to Words, Phrases, and Passages in the Dramatic Works of Shakespeare*. New York: Macmillan.
- Canby, H. S. 1947. *American Memoir* Boston: Houghton Mifflin.
- Cook, A. S. 1908. *A Concordance to the English Poems of Thomas Gray*. Boston: Houghton-Mifflin.
- Graff, G. 1987. *Professing Literature*. Chicago: University of Chicago Press.
- Howard, P. 1973. *Theses in American Literature, 1896-1971*. Ann Arbor MI: Pierian Press.
- Hyder, C. K. 1940. *A Concordance to the poems of A. E. Housman*. Lawrence KA: University of Kansas.
- Manly, J. M., and E. Rickert. 1940. *The Text of the Canterbury Tales studied on the Basis of All Known Manuscripts*. 8 vols. Chicago: University of Chicago Press.
- McNamee, L. F. 1968. *Dissertations in English and American Literature*. New York: Bowker.
- Perry, B. 1935. *And Gladly Teach*. Boston: Houghton Mifflin.
- Wright, L. B. 1931. "Handbook Learning of the Renaissance Middle Class." *Studies in Philology* 28:58-86.