Digital Paper

A Manual for Research and Writing with Library and Internet Materials

Andrew Abbott

THE UNIVERSITY OF CHICAGO PRESS
CHICAGO AND LONDON

ANDREW ABBOTT is the Gustavus F. and Ann M. Swift Distinguished Service Professor at the University of Chicago. He edits the *American Journal of Sociology* and his books include *The System of Professions*, *Department and Discipline*, *Chaos of Disciplines*, and *Time Matters*. He has twice chaired the University of Chicago's Library Board, and he played a central role in planning the university's Joe and Rika Mansueto Library.

The University of Chicago Press, Chicago 60637
The University of Chicago Press, Ltd., London

Printed in the United States of America

23 22 21 20 19 18 17 16 15 14 1 2 3 4 5

ISBN-13: 978-0-226-16764-0 (cloth)
ISBN-13: 978-0-226-16778-7 (paper)
ISBN-13: 978-0-226-16781-7 (e-book)
DOI: 10.7208/chicago/9780226167817.001.0001

Library of Congress Cataloging-in-Publication Data

Abbott, Andrew Delano, author.
Digital paper : a manual for research and writing with library and internet materials / Andrew Abbott.
pages cm — (Chicago guides to writing, editing, and publishing)
Includes bibliographical references and index.
ISBN 978-0-226-16764-0 (cloth : alk. paper) — ISBN 978-0-226-16778-7 (pbk. : alk. paper) — ISBN 978-0-226-16781-7 (e-book) 1. Report writing—Handbooks, manuals, etc. 2. Research—Handbooks, manuals, etc. 3. Searching, Bibliographical—Handbooks, manuals, etc. I. Title. II. Series: Chicago guides to writing, editing, and publishing.
LB1047.3.A22 2014
808.02—dc23

2013050782

♾ This paper meets the requirements of ANSI/NISO Z39.48–1992 (Permanence of Paper).

To Judi Nadler

in reading we aim to find not items, but arguments, descriptions, and other extended intellectual structures. We are seeking things that require thought, not simple recognition. We find them not by associating them with attractors, but by mastering their syntax or pattern. Reading is a *discursive* technique. That requires a different type of attention.

1. Modes of Reading

Like scanning/browsing/brute force, reading can be done at many levels. But in this case the level is not determined by the density of the text in terms of expected return, as it is with browsing and scanning. Rather it is determined by the particular kinds of things we want from reading. Sometimes we want simply to follow a story. Sometimes we want to be stimulated intellectually. Sometimes we want to master the bare bones of argument. Each of these has its own strategy.

A. NARRATIVE READING

Narrative is the most familiar type of reading. We read to get the story. Thus, narrative fiction is written on the assumption that you will read every word, as if the story were being told out loud. Especially in nineteenth-century fiction, this structure was exacerbated by the use of sentences so complex that they are impossible to understand without subvocalization. One must hear such a text in order to parse it.

Reading of this kind is seldom necessary in any social science research materials, primary or secondary. Few of them are written with self-conscious elegance, and most of them are conceived not as narrative but as exposition, not as story, but as analysis. As we shall see below, expository texts are amenable to various forms of accelerated reading, which take advantage of their topic sentences and section summaries to skim much of the intervening text. In many expository books, there is even a complete summary of the book's argument in the final chapter. For harried researchers, this is a useful thing indeed.

But in library research there are some occasions for narrative reading. In the humanities, of course, narrative reading is often the first method for reading primary materials. (Later readings will be not narrative but analytic.) But more generally, a library researcher reads narratively when reading "in the background." In background reading you are orienting yourself to an area, getting used to the important names and concepts, and immersing

yourself in debates and problems. You aim thereby to fill your medium-term memory with browsing attractors and to start building a general sense of your research area. To be sure, you may want to take some notes on background reading. A timeline might be helpful, if the background material is a history. Or a sketch biography, or a listing of some major arguments. But the purpose of background reading is not to master things. It is to orient yourself to the research area and to prime your browsing eyes with attractors. Narrative reading is the proper mode for this. It is a familiar technique, so I won't discuss it further.

B. MEDITATIVE READING

In most library research, only very occasionally do you want to read more slowly than everyday narrative reading. But sometimes you may want to do such "meditative reading." You read meditatively when you are looking for new general ideas: ideas about how to attack your basic project or about new theories or interpretations. So you choose a rich text, often one that is not immediately related to your specific research questions, but rather related to the general set of theoretical issues that concern you. You read every word, slowly, and allow your mind to resonate freely with the text. You attend to possible allusions, to interesting themes, sometimes even to the sounds of the words. You read a paragraph or two, then look at the wall and reflect. Then perhaps you jot a few notes in a theory folder as ideas drift into your head: "What if I simply disregarded my second theoretical question and refocused the project?" "Is institutionalist theory really helping me here?" "What about a functionalist way of organizing the subsections?"

Then you read some more. You are going to spend three or four minutes per page. Perhaps even more. You are not aiming to get through something; your aim is simply to stimulate your own thinking. You therefore let the text come apart into shreds of insight and allusion. Obviously, such a strategy works only with very rich texts. Hence meditative reading is used for important works of theory or central passages within larger texts.

In library research, you should choose for meditative reading a major theoretical work in your own project area. One of my course students was studying the rise of environmental politics in Ecuador. So I suggested reading R. G. Collingwood's *Idea of Nature*. It's a general philosophical work about the concept of nature itself. There was no need to understand Collingwood's arguments in detail; the idea was rather to let them stimulate her thinking, to point her in new directions.

Above all, meditative reading reminds us of the importance of thinking. Even a careful researcher can get caught up in the excitement of detection, the bravura of brachiation, the thrill of the chase. All the more reason to remember that research is not ultimately about discovering something. Our project does not have an answer waiting somewhere for us to arrive. Rather, we aim to assemble a new and exciting collection of found things in order to resolve an empirical puzzle and reflect theoretically about social life. The heart of research is a creative act in the researcher, not a clever detective routine. Research cannot be done without hard thinking, and meditative reading guarantees that we do that hard thinking.

One should always therefore remember that the authors we read during a research project had ideas; they didn't simply recite facts. They struggled to put those ideas into words to be printed. They said their ideas a dozen different ways because no particular way sounded quite right. They wrote long books because they thought they could tie down the ideas by specifying more about them. But all that extra writing simply made the rich ambiguity of the ideas even more evident. The ideas of an author don't become more specific as you read more of him or her. They become more generative. You become more able to anticipate, to channel an author. Only then are you able to use an author's ideas in your own thinking. To achieve this, you read meditatively.

C. SCAN READING

At the other end of the scale is scan reading. Scanning is done very quickly, in a text where most of the material is irrelevant, but some small portion is something you need and can quickly identify. You first use the index, table of contents, or headings to find the likely parts to scan. Then you rush as fast as possible through the pages selected, employing a template or search term. It helps, of course, if the latter has easily distinguishable features: capital letters, an extended pattern of several words, etc. (Thus in chapter 2, I scanned the ninety scholarly autobiographies looking for discussions of scholarly practices and within them for the word "library.") Your scanning speed is a function of various things. The less organized the text, the slower you must scan. The less obvious the template or search term, the slower you must scan. The more it matters if you miss something, the slower you must scan. The quality of the index, the detail of the headings within chapters, and—particularly—the length of the paragraphs and even of the sentences may make a big difference.

Scanning always involves a particular search term or terms. In scan reading, this template will come from the current minianalysis or data source. Because it is template-dependent, scan reading is sometimes easier with digital texts (if there is an obvious keyword). But if the template isn't specific (e.g., if you're looking for "anything about environmental issues in Ecuador"), string search won't work well. More problematic, while you can scan articles in JSTOR, you can at present scan books only through Google Books, which is prevented by copyright law from allowing full access. So you should learn how to scan-read physical books.

In scan reading you are likely to be distracted by browsing opportunities. Background reading and controlled vocabulary building will have put attractors in your head, and you will constantly be tempted to slow down to browsing speed as you see those attractors in the text. But force yourself to move on. Scan reading is a brute-force exercise: straight through, no skips, absolute focus. It will be quite tiring.

D. READING FOR MASTERY OF ARGUMENT

Mastery reading is the standard mode of full-text reading in social science. We mastery-read when we must know a text's core argument but can ignore peripheral details like examples, asides, and minor corollaries. Several variables affect the ease of mastery reading. Texts vary in length. They vary in density. Third and most important, they may or may not follow the standard rules for high-quality expository prose, which make texts easy to ransack for main points and overarching ideas. At the level of prose style, these rules include mostly short sentences mixed with occasional long ones, very clear use of referents and logical markers, short paragraphs with obvious topic sentences (usually as opening sentences), and clear logical linkage. At the book level, these rules call for clear and helpful chapter structure, strong (and possibly multilevel) subdivisions within chapters, paragraphs of sane length (two or three per page), and, above all, clear summaries at the ends of subsections, sections, chapters, and possibly even at the end of the book.

All these make a text easier to master. But even if a text is not particularly well written, you still want to apply as many of the well-written-text strategies as you can. Fall back on bad-text strategies (discussed below) only when absolutely necessary.

To mastery-read a well-written book, you first master the table of contents. Read it four or five times. Memorize it. Then scan the index to find

the dozen or so most-cited abstractions. These will be the core terms of the author's argument. Memorize these so that you will recognize them immediately as you scan pages. Your main task is to understand how these words are logically connected by the author's argument.

Third, check the last chapter to see if there is a summary of the argument. Also check the ends of each chapter for summaries. (Some writers write uninformative summaries, by the way—perorations rather than recapitulations.) If you find summaries, read them carefully; note how they combine the major terms from the index. Try to recite the book's main argument to yourself, as if you were explaining it to someone else. Remember, you are not reading narratively. You are reading "from the top down." You should always be flipping pages, always moving back and forth, always forcing yourself to summarize your knowledge so far. If you catch yourself reading steadily for five minutes, stop. You are out of mastery mode.

Once you have mastered the layout, terms, and basic argument (this will have taken about half an hour and feel like lot of work), scan the book for about fifteen minutes, looking for anything else to absorb or to add to your understanding of the argument. Then and only then, allow yourself ten or fifteen minutes to write summary notes of the argument. Refer back to the table of contents, index, summaries, and text as necessary.

This entire process should take about an hour for a well-written 300-page book. By this process, you can, for example, probably master my own first book's six theoretical chapters in about an hour. Of this hour, you will have spent as little as a quarter to a half actually "reading" sentences. The rest is all mastering the organization and finding what to read. You can be very sure that after an hour spent this way, you will know ten times as much about the book as you would if you started at the beginning and read narratively. (A really good mastery reader, for example, will have found that the first three theoretical chapters are summarized almost completely on pages 111–113.)

It is because you spend so much of this time flipping back and forth, by the way, that mastery reading is virtually impossible online. You can't flip pages fast enough or scan large blocks of text as fast. Curiously, the online environment favors narrative reading, which is useful only in a narrow range of research activities. That's why e-books—which allow paging—are becoming more common. But they're still slower than physical books.

At the article level, too, good writing is relatively easily to mastery-read. This is particularly true because there is a limited number of logical templates for articles: compare-and-contrast, case-analysis, quantitative-causal-analysis, adjudication-of-two-theories, tell-the-story, and so on. You know these templates well, and how to read them. Unfortunately, writing at the sentence and paragraph level is often pretty bad in articles, usually because of slavish imitation: sometimes imitation of scientific style, sometimes imitation of obscure philosophy or theory. Also, in an article there is no table of contents or index. So you must rely on the abstract, the headings, and the conclusion to produce your initial impression of an article.

The mastery reading of a well-written article follows the same algorithm as the reading of the well-written book. It will take about twenty minutes. First, master the abstract. Get the whole thing into your mind. Become familiar with all parts of it. Notice what aspects of it are unclear, and mark them in your mind for further investigation. You will need to read the abstract five or six times to do this. It will take five to ten minutes. Next, comparing the abstract to your own concept of an ideal version of this particular article, figure out the necessary parts of that ideal that aren't evident from the abstract.

In an empirical article, these nonevident things may be details about methods, the data description, and the major qualifications or limitations. Make a list of them in your mind as questions you need answered in your remaining time. Note that the theory and the main results of an empirical article should appear in a well-written abstract. If they aren't there, then you apply techniques given below for articles not well written. (Authors do sometimes write bad abstracts, by the way, but copyeditors generally fix them in production.)

In a theoretical article, by contrast, the entire argument will usually be outside the abstract, which will be purely formulaic. You should notice this immediately and at once scan the article for a basic summary of the main argument. There will be one somewhere. Treat this summary as if it were the abstract (read it five or six times and master it). Often, this summary will be at the end of the next-to-last section, just before the author begins to discuss "further directions." (Ignore all further directions. They are meaningless filler.) This second phase will take another five to eight minutes.

Finally, scan the text quickly for the necessary things that are missing from the abstract (or the summary or the equivalent passage in text). Then read (narratively) enough text (found by this scan) to fill the holes you have found. Then set the article aside. You are done. You should be able to process a well-written article, at least well enough to get its main points, in about twenty minutes. Write a set of short notes in summary, in a text file, just as for a book.

E. PARTIAL MASTERY READING

Most often in research, you will not need to completely master a book or article. Rather, you need to extract and master some part of it. This is partial mastery reading—in practice, the most common mode of reading in the social sciences and humanities. When we read a book or article we usually want to know only some part of it. We want to know not "what Laumann et al. said about sex" but "what Laumann et al. said about sexual identity among homosexuals"; not "what West and Zimmerman said" but "how West and Zimmerman's concept of 'doing gender' differs from Mertonian 'role performance.'" In each case we want something very specific, and we need understand only what is necessary to answer our particular questions.

Thus the central requisite for partial mastery reading is a very well specified question. As always, being a good library researcher is not about finding things but about knowing what to look for. Reading without questions always reverts to narrative reading, and narrative reading always fails for anything but the first reading of a novel. (Or sometimes of a history. In practice, however, narrative reading fails even for history most of the time.) Like mastery reading, partial mastery reading is very hard work. If you do not have to rest every half hour because of sheer reading exhaustion, you are not working hard enough.

In partial mastery, you begin by specifying the things you must retrieve from the text: Results? Theory? Steps of an argument? Data? Data elicitation methods? Statistical procedures? Attitude to this or that concept? Text? Scholar? Finding? Usually it will just be one or two of these things. Then tell yourself "This means that I do not need to know. . . ." and recite to yourself the things you can and must ignore. (If you are a good browser, the temptation to slow down is always great.) Then read the text rapidly and mercilessly, insistently asking yourself, "Is what I am looking for on this page?" If it is not, then immediately go on. Do not look at this clever paragraph or that interesting citation. Just go on. You should be able to "partially master" any article in about ten minutes. (Not ten seconds, by the way; take the time to find what you really need.) Again, this will be very hard work. Do not underline or select-and-paste. That takes extra time and is postponing mastery of the text. Rather, extract what you want the first time and write your own note about it in a text file with the citation. You must aim to think the ideas of the text, not simply to find them.

You will note that my recommendations for mastery reading always involve two stages: (a) figure out what you want to discover, and (b) search for

that and that alone. The most common mistake in research reading (also in scholarly reading more generally) is to read narratively when you should be reading for mastery. Never read a text without having specified ahead of time what you want get from it. Any other approach degenerates into narrative reading. Save that for Tolstoy.

TECHNIQUES FOR WORK THAT IS NOT WELL WRITTEN VARY, FOR THERE are different types of hard-to-read material, and different reading strategies are appropriate for each. First, there is material that follows a prose aesthetic different from that of modern English exposition. In the current research environment, one encounters such materials mainly in two places: in foreign language theory and in parts of the humanities. Translated foreign language theory (Foucault, Bourdieu, Habermas, et al.) will be read mostly for theoretical stimulation (in meditative mode). It is fine for that purpose. Modern French theoretical prose is best read very quickly, without any attention to logic and argument, of which there is often very little. But it is usually beautiful and always stimulating. German theoretical prose has the somewhat different problem that it usually employs specialized terminology from the German idealist heritage and makes little sense unless you already know some Kant, Hegel, Heidegger, and so on. If your research project requires reading German theory, find a good trot.

As for humanistic writing, it too can be very useful for meditative use in library research. Occasionally it is useful for partial mastery, particularly the older material that is often highly focused and articulate. In more recent years, however, many articles in literary studies affect a deliberately obscure prose. Often, their aim is to be art or culture as much as it is to discuss art or culture. Such articles self-consciously demand narrative reading. Occasionally, they are worth it.

As for badly written social science, of which there is a very great deal, you will have trouble reading it as efficiently as you can read well-written material. (Do not let this make you think it more important.) Since the article or book itself is not well organized, you yourself have to hypothesize and complete the necessary organization. You begin by deciding on the closest template to what you are reading: Is it a story? (Perhaps put better, is it trying to be a story?) Is it a case study? An adjudication of theories? Once you have decided on which of these it is, imagine the ingredients that are necessary for a good version of that template. Scan the article or book to find as many of them as you can, then build in your mind a version of the article/book as

it should have been written. Do not allow yourself any more time than for a well-written work; an hour for a book, twenty minutes for an article. Most likely, the weak writing is a sign of careless thinking and there isn't anything wonderful to find, no matter how long you work with the text.

THE VARIOUS MODES OF READING PROVIDE A DIVERSIFIED REPERTOIRE for dealing with the various secondary materials you confront in a library research project. You will use narrative reading on the background materials that you read to orient yourself at the outset. You will also use it when new areas and subareas emerge during the research. You will read meditatively when you need to see the project from the outside: when you need to rethink, perhaps redesign. You will scan-read throughout the project, because most secondary sources will be useful only in part; you need to extract what they say and move on. You will also find yourself scan-reading some of your primary sources as you try to uncover which of them are the most worthwhile, the most productive in terms of your research questions.

Mastery reading is necessary for the theoretical and empirical works central to your project. Even here, however, you do not need to master every sentence, every subargument. You need rather to have a firm command of the major arguments, the major ideas, the major empirical analyses. Remember, the center of your project is your work, not someone else's. Partial mastery will suffice for most of the dozens of peripheral secondary sources that may bear on your project. Get what you need and move on at once.

All of these, however, are modes of reading mainly for secondary materials. That is, they are ways of reading work that is itself self-consciously outside the subject of your research itself. As for primary materials, they require a quite different approach.

2. Primary Materials

A. NOTE-TAKING

First of all, primary materials require disciplined note-taking. Sometimes this is physically necessary: you may be reading primary materials in an archive that does not permit you to photograph them with a digital camera or to otherwise copy them. But you also will take notes as part of analytic reading. You must abstract the themes and issues from primary materials so they are available to you instantly once you start analysis. Note-taking is as much an analytic strategy as it is an occasional necessity.

Some rules, then, about note-taking. Do not keep notes on primary materials in continuous text. Keep them in bits so they can later be indexed and possibly put in different places. These bits should be relatively short. If you keep notes in thirty-line paragraphs, you may have to distribute copies of a given paragraph to four or five different analysis folders (physical or virtual). Then you have to underline, or otherwise demarcate, those parts of the paragraph relevant to that folder, and you will have unnecessary, peripheral materials (the nonunderlined bits) associated with them.

It is better to take your primary-material notes in shorter, codable units. The note-taking unit should be small enough to have only one or two analytic uses. Typically, this will be two to ten lines of text. Then make a line break (in straight text notes) or start a new indexable unit (if you are working in a codable database). In the database approach, this level of detail seems to me cumbersome, as you will end up with long units with several tags, rather than putting the vast majority of units in unique locations. (I myself take notes on primary material in the form of straight text with blank-line breaks to mark new units.)

Whatever the size of unit you choose, it must carry with it complete citation information. Get into the habit of typing some identifier at the beginning of each note-unit. An example from my archival notes: "2.19, Ltr EOL to CEB, 24 Feb 72 says that he has doubts about the managing editor." This telegraphic note means that the content comes from box 2, folder 19. It is a letter from Edward O. Laumann to Charles E. Bidwell, dated 24 February 1972, complaining about the managing editor of the journal in whose records this letter appears (the *American Journal of Sociology*). Since the entire string of notes comes from the *AJS* archival files, I don't bother to include that information here.

This telegraphic form of citation explains why you must build a detailed list of people (or organizations, or whatever are your main units of analysis) in your data folders. It will contain the abbreviations of names (of people, institutions, organizations, etc.) that figure in these short identification headers for primary material.

B. DOUBT

Reading of primary materials is not like reading of secondary materials. Secondary materials are part of a scholarly literature. They have been vetted, edited, challenged. By contrast, primary materials stand alone. You may be their only reader. So your responsibility is greater. You must constantly ask critical questions of any text: Who wrote or produced it? For what purpose