2015-11

Reading in the 21st century; reading at scale

Ammerman, Jack

http://hdl.handle.net/2144/22850

Boston University
... It is rather when
We gloriously forget ourselves, and plunge
Soul-forward, headlong, into a book's profound,
Impassioned for its beauty and salt of truth—
'Tis then we get the right good from a book. (Browning [18-?], 24)

Speaking in 2003 about the future of books, Umberto Eco recounted Victor Hugo’s story of a 15th century priest, Claude Frollo, who looking first at the towers of Notre Dame de Paris and then at a printed book on his table “whispers ‘ceci tuera cela’: this will kill that, or, in other words, the book will kill the cathedral, the alphabet will kill images. The book will distract people from their most important values, encouraging unnecessary information, free interpretation of the Scriptures, insane curiosity” (Eco 2003, 8). Before the invention of printing, manuscripts were accessible only to an elite group of literate persons. The illiterate masses learned the stories of the Bible, the saints, the life of Christ, and even moral principles through the images of the cathedral. Now they would learn to read.

Frollo was correct. Books emerged as primary carriers not only of information, but of culture. Reading both informs our most important values and distracts us from them. Reading is a process by which we attempt to discern the most necessary information all the while awash in a sea of unnecessary information. The democratizing impact of reading extends far beyond the Scriptures. And, reading has become both a stimulus of and tool for satisfying our “insane curiosity.”
Why read?

Why do we read? We could be doing other things with our time: sports, leisure, sleep, hobbies, household chores. For centuries, readers have provided answers that generally fall into two categories, those that identify reading as a means to an end and those that identify reading as an end in itself. For some, reading is a formative practice. “Ultimately we read – as Bacon, Johnson, and Emerson agree – in order to strengthen the self, and to learn its authentic interests” (Bloom 2000, 22). Others claim we read more instrumentally, to accomplish something beyond ourselves. “[W]e read with some purpose, in which literature is a means (to knowledge, escape, a particular sensation, entertainment, or even attainment of a kind of cultural authority)” (Kriner 2014, 3). These two categories of reading tend to collapse together, or at least represent more of a continuum than binary options. Efforts to distinguish the two are less important than affirming the importance of reading, particularly long-form reading, to learning and culture. The information environment in which we read today is greatly shaped by digital media and technologies. In the following paragraphs, I will explore the impact of digital media on our practices of reading in the 21st century. First, it is important to understand these emerging practices of reading in the context of the history of reading.

Close Reading

When we talk about reading, we often assume linear reading of a long-form text. Expanding on this, educators describe ‘close reading’ or ‘deep reading’ as “reading to uncover layers of meaning that lead to deep comprehension” (Boyles 2013). Through close reading, the reader engages the text critically, discovering details and patterns that allow the reader to
develop a precise understanding of the text’s form and meanings. This, model often considered to be the gold standard by literary scholars, is thought to be threatened by some.

In the last 25 years, our culture’s relationship to books, reading and what some call liberal education has changed radically with the proliferation of high-speed networks, mobile computing and communication devices, and digital content. In 1994, Sven Birkerts lamented, “The printed word is part of a vestigial order that we are moving away from – by choice and by societal compulsion” (Birkerts 1994, 118). The National Endowment for the Arts (NEA) released a report in which they identify the decline in the reading of literature. “Reading at Risk presents a distressing but objective overview of national trends. The accelerating declines in literary reading among all demographic groups of American adults indicate an imminent cultural crisis” (National Endowment for the Arts. 2004, xiii).(Ivanov 2009, xiii) The NEA report focused on literary reading, but other surveys (Digital Book World 2014);(United States Department of Labor. Bureau of Labor Statistics 2015) report similar declines in reading defined more broadly.

Such claims raise the question: Are we merely in the midst of another ‘this-kills-that’ story? Frollo’s Cathedral remains, and there is nothing to suggest the imminent demise of print books despite rapid growth in the number of eBooks. Or are such crisis discourses obscuring more fundamental changes in the role of books and reading in our culture? Regardless the answer, most will agree that our bookish culture is shifting, and with it the way we read.

Jay David Bolter introduced the phrase, “the late age of print,” as a way of describing the “transformation of our social and cultural attitudes toward, and uses of” print technologies (Bolter 2001, 3). Bolter is not writing an elegy for print books. By ‘late,’ he suggests that something has changed. Print books continue to shape how we think, act, and communicate.
Yet they do so in an information landscape that is dense with digital media and technologies. The idea of the book remains the same, yet is somehow changed. And with that change, we should expect changes in the way we read.

Reading has a long history, of which the image portrayed by Elizabeth Barrett Browning above and embraced by the NEA is but a part. Far from a simple means of converting symbolic letter forms into meaningful information, “it is a complex set of physical, cognitive, and social practices that have varied with time and place” (Levy 2001, 107). Historians of reading often describe the “interplay of binary opposites: reading by turning the leaves of a codex as opposed to reading by unrolling a volumen, reading printed texts in contrast to reading manuscripts, silent reading as distinct from reading aloud, reading alone rather than reading in groups, reading extensively by racing through different kinds of material vs. reading intensively by perusing a few books many times” (Darnton 2009, 172). Indeed, in the monastic practice of reading described by Illich, reading was not just aloud, it was a means of embodying the page. Reading was more of a “carnal activity: the reader understands the lines by moving to their beat, remembers them by capturing their rhythm, and thinks of them in terms of putting them into his mouth and chewing. No wonder that pre-university monasteries are described to us in various sources as the dwelling places of mumblers and munchers” (Illich 1993, 54).

Naturally, monastic reading would have focused on the Bible and a small number of primarily religious texts that were read repeatedly as a sacred practice. Such intensive reading was practiced even outside monastic communities. Classic texts like Aristotle, Cicero, and the Bible were considered core texts for students to master in schools and universities. The texts were read and re-read. Memory devices were used to remember them and notes and marginal
annotations were written as students worked to make additional connections with each re-
reading (Burke 2000, 180).

We imagine that our practice of reading a book in a linear fashion, following the flow of
the narrative from beginning to end has always been the model for reading. Yet, early modern
Englishmen frequently read multiple books concurrently, jumping from book to book. “They
broke texts into fragments and assembled them into new patterns by transcribing them into
different sections of their notebooks” (Darnton 2009, 149). ‘Commonplacing,’ as it was known,
has a long history, dating from the 12th century and continuing well into the Victorian era.
Excerpts from books being read were copied and arranged thematically to represent the
reader’s synthesis of what he or she had read. Commonplace books were standard tools used
by readers as a way of construing the world and were especially effective for organizing what
seemed a rapidly growing body of information.

The shift from intensive to extensive reading coincided with the desacralization of the
book (Darnton 2009, 203) (Burke 2000, 179). Books became more plentiful and began to be
treated more like commodities. Reasons for reading began shifting from spiritual practice to
efforts to understand the world. And the established patterns for reading and mastering a small
number of texts were less adequate to enable reading a rapidly growing number of books.

The increasing rate of publication prompted changes in how books were published as well
as in the types of materials that were published. Indexes, tables of contents, and lists began to
be added to books being published to help navigation as readers began reading in non-linear
fashions. Dictionaries, encyclopedias, and other “reference” books not intended to be read
intensively were published. The need to organize and access a growing body of information began to shape not just the practices of reading but of writing as well (Burke 2000, 183).

The rate of book publication has only increased. An analysis of the books in the OCLC WorldCat database, revealed that the rate of books published each year increases in what resembles an exponential curve. “Approximately half of all books held in the system-wide collection were published after 1977” (Schonfeld and Lavoie 2006). A smaller study of novels published 1969-2010 reveals a similarly rapid increase in the rate of publication (Algee-Hewitt and McGurl 2015, 2). An analysis of the publication dates for publications held by the Pitts Theology Library at Emory University, one of the largest collections of theological books, reveals a similar curve, though with a little more fluctuation. The titles in this collection reflect not only the rate of publication, but acquisition patterns. Nonetheless, the rapidly increasing rate of publications appears to be correlated with not only what, but the way we read.

The rate of publication is only one factor influencing modern reading practices. The increased availability of eBooks and eBook readers has become a significant driver of change. The Pew Research Center’s Internet & American Life Project has been documenting some of those changes. In 2012, the Project reported “In the past year, the number of those who read e-books increased from 16% of all Americans ages 16 and older to 23%. At the same time, the number of those who read printed books in the previous 12 months fell from 72% of the population ages 16 and older to 67%.” A fifth of all American adults claimed to have read an e-book in the past 12 months. And notably, the average e-book reader read more books across all formats (mean = 24) as compared to readers of print only books (mean = 15). “Fully 42% of readers of e-books said they are reading more now that long-form reading material is available
in digital format” (Rainie et al. 2012). The percentage of adults reading e-books continued to increase, 28% in 2014 up from 23% in 2012. And in that 2014 study, over 50% reported having “a dedicated handheld device—either a tablet computer like an iPad, or an e-reader such as a Kindle or Nook—for reading e-content” (Zickuhr and Rainie 2014).

Electronic formats are not always desired, however. E-books are clearly preferred for rapid access and portability, particularly when traveling or commuting. Print formats are strongly preferred when reading to others, particularly children. When sharing a book with others, print is preferred. Even when an e-book is preferred, it is not always available, or easily discoverable. 23% of Americans who consume e-content “say they find the material they are seeking “only sometimes,” “hardly ever,” or never available in the format they want” (Rainie et al. 2012).

Clearly the affordances of e-books and print books influence preference, though preference is often situational. It is not simply about convenience, access, or the sensual experience of reading in one format versus the other. The same reader might prefer one format or the other depending on a variety of factors including the length of the text, the type of book, and the task at hand. Critics of e-books frequently express concern beyond these obvious differences, though. They raise concerns about whether close reading is even possible with e-books.

Hyper-reading

Research indicates that we read text onscreen differently than in print. Reading web pages in a browser is a “rapidly interactive activity. Even new pages with plentiful information and many links are regularly viewed only for a brief period” (Weinreich et al. 2008, 5:27). The
Nielsen Norman Group reports that when reading a page on a website, users eyes move rapidly across the words in a pattern that looks like the letter ‘F.’ First the users’ eyes scan horizontally across the top, then they move mid-way down the page and scan horizontally again, though usually covering a shorter area than the top scan. The final scan is a vertical scan down the left side (Nielsen 2006). Consequently, readers have time to read at most 28% of the words on the web page (Nielsen 2008).

James Sosnoski introduced the notion of ‘hyper-reading’ that is characterized by filtering (such as searching on Google), skimming, pecking, imposing, filming, trespassing, de-authorizing, and fragmenting (1999, 163). We might update Sosnoski’s description by adding ‘juxtaposing’ as when multiple windows or screens are open and ‘scanning’ (Hayles 2012, 61). This type of reading pre-dates online reading, though. Scholars have long used a wide range of reading techniques including scanning for significant content in the midst of an overwhelming amount of less useful information (Guillory 2008, 12–13). In an information rich environment, information has become abundant, and attention has become the scarce resource. Even in print environments, scholars develop strategies to quickly eliminate extraneous information to enable focus of attention on what is deemed important.

Interesting research on the impact of digital reading devices on reading time and comprehension seems to indicate that hyper-reading may be less problematic when using dedicated reading devises such as the Kindle 3 eBook reader.

Results indicated that those reading printed materials had faster reading times than those reading from eBook readers and tablets. Participants found the tablet the most usable, followed by the eBook reader, and the printed material was considered the least usable. There was no effect of text presentation format on reading comprehension, supporting the use of eBook readers and tablet computers in academic environments. (Connell, Bayliss, and Farmer 2012)
Even if we are able to use hardware and software to address reading comprehension, some are concerned about the impact of reading online for the way our brains process information. Nicholas Carr’s book, *The Shallows: What the Internet is Doing to our Brains*, reflects his own awareness of cognitive shifts.

As McLuhan suggested, media aren’t just channels of information. They supply the stuff of thought, but they also shape the process of thought. And what the Net seems to be doing is chipping away my capacity for concentration and contemplation. Whether I’m online or not, my mind now expects to take in information the way the Net distributes it: in a swiftly moving stream of particles. Once I was a scuba diver in a sea of words. Now I zip along the surface like a guy on a Jet Ski. (2010, 6–7)

Hayles cautions that while Carr is generally conscientious in reporting research results, he does occasionally state the results from these research reports more strongly than the original reports in order to support his view (2012, 69). She does suggest, though, that associated with hyper-reading, we are witnessing a shift in cognitive styles that can be seen in the contrast between deep attention and hyper attention. Deep attention, the cognitive style traditionally associated with the humanities, is characterized by concentrating on a single object for long periods (say, a novel by Dickens), ignoring outside stimuli while so engaged, preferring a single information stream, and having a high tolerance for long focus times. Hyper attention is characterized by switching focus rapidly among different tasks, preferring multiple information streams, seeking a high level of stimulation, and having a low tolerance for boredom. (2007, 187)

Hyper-reading and hyper-attention are effective and appropriate techniques for discovering, organizing and accessing information in a media intensive environment. The challenge is to hold these techniques in balance with deep reading and deep attention. The hardware and software used for reading may aid in maintaining this balance. The task remains, however, to develop strategies to nurture both deep and hyper-reading.
Deep reading and hyper-reading are joined by a third methodology for reading that benefits from mass digitization efforts such as the Google Book Index and Internet Archive along with smaller efforts mounted by libraries around the globe. The creation of textual corpora from these digitized texts invites the use of computers to analyze them as data.

**Distant Reading**

Sometimes called machine reading, machine learning or distant reading, Hayles distinguishes this new model of reading from hyper-reading by suggesting that hyper-reading is computer-assisted human reading. The introduction of computational tools for textual analysis she calls human-assisted computer reading (2012, 70). The most simple of these tools use a variety of algorithms to accomplish word-frequency counts that can then be visualized on timelines like Bookworm (bookworm.culturomics.org) or Google Books Ngram Viewer (https://books.google.com/ngrams), create word-clouds. More sophisticated programs have been developed to compare phrases, do topic modeling, identify genres, and do probabilistic reasoning. These approaches typically identify and analyze patterns across a textual corpus, often much larger than an individual reader could read using deep reading methods.

Though we might assume that technology is driving the adoption of these tools, scale is the most important issue (Hayles 2012, 27). The number of books any one person can read unaided by computers is actually quite small. Even if a person read one book per day from age fifteen to age eighty-five, she would read little more than 25,000 books. Using repositories like HathiTrust (www.hathitrust.org), one can easily run algorithms against millions of books.

The importance of scale is not simply the ability to include more books, but to broaden the context of one’s investigation. When using a deep reading methodology, we typically work
with a small canon of books that are recognized to be the most important works to read for our investigation. Excluded are those works that might be considered ‘ordinary.’ Our small canon may include the most important works, but we have lost the broader context in which those ‘important’ works exist. “Scale changes not only the quantities of texts that can be interrogated but also the contexts and contents of the questions” (Hayles 2012, 28)

For example, in the last 100 years, the purpose of education, and I would suggest reading, has become much more utilitarian. Dan Berrett (2015) claims that Governor Ronald Reagan’s 1967 statement that "we do believe that there are certain intellectual luxuries that perhaps we could do without," (Reagan 1967, 6) was a tipping point between two visions of education. “Learning for learning’s sake might be nice, but the rest of us shouldn’t have to pay for it. A higher education should prepare students for jobs.” (Berrett 2015)

While Reagan’s statement may have been highly visible, an analysis of books published from 1800 through 2000 indicates the tipping point probably occurred much earlier in the 20th century. Figure 1 plots the number of words per million in books published since 1800 for the words, ‘read’ and ‘education’ in the HathiTrust repository. While the frequency of the word ‘read’ was relatively stable throughout the 19th century, the frequency of the word ‘education’ rapidly surpassed it in 1911. In the 20th century, the frequency of the word ‘read’ peaks in 1927 and begins a gradual decline from which it hasn’t recovered.
Overlaying the plots for these two words with a plot for the word ‘employment,’ (Figure 2) reveals that the frequency of use for ‘employment’ closely parallels the frequency for the term ‘education,’ suggesting a utilitarian shift in expectations for education.

Clearly these two word-frequency plots are insufficient to make conclusive statements about the changing practices of reading or education much less causality. The introduction of computer analysis of the large corpus does, however, allow us to gain a different perspective on and perhaps better understanding of the context of the phenomenon being considered. By
means of the use of computational tools of textual analysis, the reader is able to ask questions that were previously impossible to answer, if the questions had even been imagined.

Though these computationally intensive approaches to reading have received much recent attention, most digital humanists trace their roots to Roberto Bussa’s late 1940s project to produce an automated concordance to the works of Thomas Aquinas using a computer. An Italian Jesuit priest, Father Bussa created a “radically transformed, reordered, disassembled, and reassembled version” (Ramsay 2011, 1) of Aquinas’ works. Even though we are accustomed to such works, we have usually considered the products (concordances, etc.) to be tools to assist our reading. We have largely ignored the process of their creation. To consider these computational processes as ‘reading’ expands our understanding of the word beyond simple parsing and comprehension of words to comprehension of a text or even a corpus. And with this comprehension we utilize a set of tools enabling the reader to move beyond comprehension to critical engagement with the text.

Franco Moretti, provocatively advocates for distant reading:

At bottom, it’s a theological exercise—very solemn treatment of very few texts taken very seriously—whereas what we really need is a little pact with the devil: we know how to read texts, now let’s learn how not to read them. Distant reading: where distance, let me repeat it, is a condition of knowledge: it allows you to focus on units that are much smaller or much larger than the text: devices, themes, tropes—or genres and systems. And if, between the very small and the very large, the text itself disappears, well, it is one of those cases when one can justifiably say, Less is more. If we want to understand the system in its entirety, we must accept losing something. We always pay a price for theoretical knowledge: reality is infinitely rich; concepts are abstract, are poor. But it’s precisely this ‘poverty’ that makes it possible to handle them, and therefore to know. (2000)

Here Moretti refers to a specific kind of knowledge gained by reading from a macro level perspective. From a distance, the reader is able to see connections and patterns, forms and
structures that are not apparent on ‘close’ examination. This is not a rejection of close reading. It is rather a growing awareness of a kind of parallel universe in which the primary texts of his investigation have become data. Even in this parallel universe, human interpretation remains essential to critical reading.

If one can read at a distance as Moretti and Hayles suggest, it is not a great leap for Stephen Ramsay to propose an “algorithmic criticism” that attempts to “locate a hermeneutics at the boundary between mechanism and theory.... [to] create tools—practical, instrumental, verifiable mechanisms—that enable critical engagement, interpretation, conversation, and contemplation” (2011, x). If there is to be an algorithmic criticism, it will be both revolutionary and conservative, both new and old. Computation, suggests Ramsay, doesn’t imply an alternative hermeneutical procedure but reimagines existing hermeneutical procedures at new scales, new speeds, and new sets of conditions (2011, 31).

While there is no scientific evidence of the impact of distant reading on the way we think comparable to the evidence available for hyper-reading, it would be an interesting comparison. While algorithmic criticism may simply be a set of tools to exploit the increasing availability of digital texts, it may better be seen as a methodology for coping with them like hyper-reading is a strategy for coping with the amount of information available online. Ramsay suggests “in the end, it is simply an attitude toward the relationship between mechanism and meaning that is expansive enough to imagine building as a form of thinking” (2011, 85). If Ramsay is correct, the acts of textual encoding, analysis, and building web sites, are examples not only of thinking, but of embodying the texts in new ways. Just as Illich described monastic reading as a “carnal activity” (1993, 54) in which the readers were embodying the text through their physical
movements and by reading aloud, distant reading and algorithmic criticism may be seen as new ways of embodying the text through building.

**How Read?**

Earlier, I posed the question: Why Read? Clearly we read for many reasons, among them to make sense of what seems to be a rapidly expanding and ever more complex world. Perhaps the more important question is: How do we read? In the past ten years, “cloud computing” has come of age as companies have worked to develop flexible and scalable solutions to their technology needs. In the past few years, Lorcan Dempsey, Vice President, OCLC Research and Chief Strategist, focused the attention of libraries on what he calls “rightscaling” (Dempsey 2013) as they develop solutions to address their needs. Some issues are appropriate to address on a local scale, while for others, it only makes sense to address them on much more collaborative regional or even global scale.

As we have seen, issues of scale have always been significant factors in determining reading methodologies. Ultimately, no single methodology has replaced all others. Deep reading remains significant even as hyper-reading and distant reading methodologies continue to develop. The more significant issue is rightscaling, selecting the best methodology for the task at hand.

**Works Cited**


