"It may change my understanding of the field": Understanding reading tools for scholars and professional readers

Ray Siemens, Cara Leitch, Analisa Blake, Karin Armstrong, & John Willinsky

2009

DHQ: Digital Humanities Quarterly
2009
Volume 3 Number 4

"It May Change My Understanding of the Field": Understanding Reading Tools for Scholars and Professional Readers

Ray Siemens <siemens_at_uvic_dot_ca>, University of Victoria
Cara Leitch <cmleitch_at_uvic_dot_ca>, University of Victoria
Analisa Blake , University of Victoria
Karin Armstrong <karindar_at_uvic_dot_ca>, University of Victoria
John Willinsky <john_dot_willinsky_at_ubc_dot_ca>, University of British Columbia/Stanford

Abstract

As the amount of scholarly material published in digital form increases, there is growing pressure on content producers to identify the needs of expert readers and to create online tools that satisfy their requirements. Based on the results of a study conducted by the Public Knowledge Project and introduced at Digital Humanities 2006 (Siemens, Willinsky and Blake), continued and augmented since, this paper discusses the reactions of Humanities Computing scholars and graduate students to using a set of online reading tools. The results of our study reveal both the potential strengths and perceived weaknesses of online reading environments. Understanding how users read and evaluate research materials, anticipating users' expectations of the reading tools and resources, and addressing user concerns about the availability of online material will lead to improvements in the design and features of online publishing.

Introduction

In 2001, the Public Knowledge Project at the University of British Columbia began work on a series of "reading tools" under the direction of John Willinsky. These tools were intended to provide readers of online scholarly journal articles with access to relevant related materials. The reading tools (originally called Research Support Tools) were incorporated into the open source software designed to manage and publish journals being developed by the Public Knowledge Project. In concept, the tools draw on the sets of links that accompany the journal articles made available online by Highwire Press, PubMed, ScienceDirect, and other online journal portals. These links enable readers to find related studies and other papers by the authors, typically from within the portal's database. 1

In developing these tools, however, the goals of the Public Knowledge Project are somewhat broader than those of other journal sites. This follows directly from the project's interest in supporting *open access* to research and scholarship through the development of open source software systems. Open access, after all, implies *public* access to a body of work. The Public Knowledge Project's tools were intended not only to support scholars reading in their area of expertise, but also to support relatively inexperienced readers of research, including students, readers from other professions, and members of the public. For that reason, these tools were designed to draw on open access sources of scholarly work, as well as to connect the user to media, government, and other non-scholarly sources. 2

This study examines the use of these reading tools by thirteen humanities computing scholars and graduate students. Our inquiry focuses on the degree to which, and the ways in which, such tools can contribute to the reading practices of domain-experts.[1] In the course of the study, readers were asked about the value of readily being able to check related studies not cited in the article; to examine work that has followed on the study reported in the article; to consider additional work that has been done by the same author; and to consult additional sources on the topic outside of the academic literature. The choice of humanities computing was based on the tendency of this field to bring together scholars with a strong understanding of critical textual practices as part of a humanities tradition along with a strong interest in new text-related systems. Those working in humanities computing have, for example, been active in creating a number of digital tools for textual analysis (e.g., TaPOR) and other forms of research and scholarly work [Schreibman, Siemens & Unsworth 2004]. This involvement in the conceptualization, design, and creation of reading tools is one of the factors that distinguishes humanities computing scholars from their colleagues. Like all scholars in the humanities, domain-expert readers in humanities computing are concerned with acquiring, assessing, and building on existing research material. Where Humanities Computing scholars differ, however, is in their tendency to evaluate their own research and reading habits alongside their research material. Every act of reading becomes another opportunity for the researcher to refine his or her understanding of textual practices. 3

Our interviews with domain-expert readers revealed that reading tools can make a definite, if limited, contribution to readers' critical engagement with journal articles. Participants made it clear that the tools' abilities to bring well-established, reputable scholarly sites within reach was an important benefit, given their apprehensions about the quality of materials available online. While the tools helped readers to position an article and its author in relation to other work in the field, participants felt the tools did little to further their comprehension or understanding of the article itself. This was largely because these readers already have critical methods and sources in place, which the tools were emulating with less-than-perfect accuracy or precision. These readers did see as much, if not more, value in these tools for their students in terms of helping them to evaluate an article and position it within a larger field. In sum, participants made it clear that reading tools could contribute to their reading of an academic article. The degree of usefulness of the tools depends on how well the tools can assist them in (a) evaluating and placing an author's work within the larger field, (b) gaining a sense of that field from the online discussion of issues, (c) organizing the retention of the relevant research materials they encountered, and (d) taking full advantage of the library resources to which these readers have access. 4

Prior Studies

The reading practices of expert readers have been the focus of a number of prior studies. The results of this study speak to aspects of domain-expert reading that have not been previously addressed. In particular, our analysis has been motivated by Sam Wineburg's claim that "to understand the 'historic sense', we must study people as they engage in the process of historical enquiry" [Wineburg 1991, 73]. The act of using the reading tools seemed to prompt participants to think about how they usually did things. During the course of our study, we found that participants spent as much time reporting about their own reading practices as they did commenting on the reading tools. The current literature largely focuses on comprehension strategies at the level of the article rather than on how readers develop their understanding of a field by positioning articles and authors within that field. For example, in a study with close parallels to this one, Wyatt et al. studied fifteen social science professors who, on choosing a journal article to read, provided a commentary on what they felt was important about their approach to reading the work [Wyatt et al. 1993]. These readers drew upon a rich body of background knowledge to assist them in comprehending the article. The researchers found that these readers also had reading strategies that involved a continual process of anticipation and prediction. Their reading process called for self-regulation and self-monitoring as they jumped back and forth within the text in search of particular pieces of information, formed conclusions, evaluated research methods, assessed the quality of writing, and weighed the relevance of the topic. In this way, these readers tackled an article's main ideas and tested the adequacy of its claims as part of a largely self-contained system. 5

In a study that compared how historians and students read, Sam Wineburg notes that scholars possess "knowledge of how to establish warrant and determine the validity of competing truth claims in a discipline" [Wineburg 1991, 83]. Wineburg also uses the example of a historian reading outside her own area of historical specialization to demonstrate that "her expertise lay not in what she knew, but in what she was able to do when she did not know" [Wineburg 1991, 83]. The sense is that these readers focused on utilizing reading strategies that determine the validity of the content in the text they are reading. In a later piece, Wineburg discusses how two historians engage in "the active processes of creating historical contexts" to further their understanding of the history at issue, while the study reported here reveals an additional concern among domain-experts in actively establishing the disciplinary context within which to place what they are reading [Wineburg 1998, 339]. 6

Pressley and Afflerbach report findings of particular relevance to this study, namely that skilled readers are "opportunistic" in the sense of being prepared to exploit all available textual clues (of which the reading tools might be thought to provide an additional set) while seeking to identify consistencies and inconsistencies between the text being read and their prior knowledge (which might be said to involve adjusting their evaluation of an author's position within a field). In a similar vein, Patricia A. Alexander sums up the result of a number of studies of "highly competent learners" as demonstrating that these learners engage in a goal-directed behavior that is personally satisfying by bringing "a body of relevant prior knowledge" matched by "a rich repertoire of both surface-level and deep processing strategies," again focused on comprehending the text being read. To this, our study would add the satisfaction gained by readers when they increase their hold on the larger field and those working within it. This focus on the comprehension of content from within the text has to do with the goals of this research, which involves helping novice readers improve their comprehension in school disciplines. It is often assumed that learning who is doing what within a field is of less importance to students than to expert readers. That may be an assumption worthy of further examination. 7

Reading Tool Design

The reading tools developed by the Public Knowledge Project sit in a frame to the right of an article or a paper after it has been published using Open Journal Systems or Open Conference Systems. (The longer-term vision for the reading tools is that they would be able to serve other systems, as a browser plug-in or part of a library's portal.) The tools are organized into two sets. The top set of tools provides information about the article itself: its authors, indexing information, the sponsors of the research, supplementary files accompanying the article, and definitions of unfamiliar words. There are also email links to contact the author or notify a colleague, as well as the ability to leave a comment on the article for other readers to consider. 8

Figure 1. An article abstract with reading tools listed to the right.

The second set of tools, which appears below the first set, provides access to related items drawn principally from open access databases that do not require readers to belong to a particular research institution. These tools, which are subject specific in most cases, have been created in 19 disciplinary fields, from agriculture to social sciences, and have been designed so that journal editors can select the appropriate set for their journals. Each of the tools in this second set is made up of relevant searchable databases.[2] If the editor has selected the Humanities reading tools set, the Related Studies tool allows the reader to search The English Server, Voice of the Shuttle, NetSERF: The Internet Connection for Medieval Resources, and five other relevant databases.[3] 9

The second set of reading tools also includes tools that reach outside the scholarly community. These tools connect readers to government databases, newspapers and media sites, as well as to teaching sites aimed at a broad range of educational levels. These tools also provide a broader context for all readers, both lay and expert, by including material such as media coverage of a topic or subsequent development of related government policies. Although this material is directly applicable to a limited range of research, it might be of interest to researchers working in related areas. 10

Another crucial factor in delivering relevant materials to readers is the accuracy and suitability of the search terms used by the reading tools. The search terms, which appear when the reader clicks on a reading tool, are based on the first two keywords provided by the author of the article. Readers can edit the terms, as well as substitute search terms that focus on a specific aspect of the article or a single concept within it. Still, the author's keywords provide a strong starting point not only for the search, but also in identifying the intended thrust and focus of the work. As might be expected, the search conducted by the Author's Work tool uses the name of the author (or authors), while the Definition tool uses any word that the reader selects in the body of the article. 11

What these reading tools contribute to this new reading environment is an extension of the traditional *paratexts* that domain-experts are so adept at reading, such as the title page, copyright information, table of contents, footnotes, indexes, appendices, and references [Stanitsek 2005]. The paratexts of print have long contributed to strengthening the internal coherence of a text while also connecting the text to a greater body of work by identifying its position within a field. In the online world, one review of hypermedia studies suggests that the domain-experts' greater command of the content means that they are interested in "the location of detailed information related to specific entities," while non-expert readers are in need of a "global picture of the material" [Chen, Jing-Ping & Macredie 2006, 262]. This study suggests that, for domain-expert readers, the detailed information provided by reading tools is directed toward a more global picture of the field. 12

Method

This study of the reading tools is concerned with their potential value to domain-expert readers in the field of Humanities computing. In designing the study, our goal was to recruit between 12–15 participants. Because of the small community within Humanities computing, several colleagues of the investigators were contacted via email to participate in the study. If an individual did not respond, a follow up email was sent. Fourteen Humanities academics from U Victoria and U British Columbia were eventually selected to participate in the study, ten during the initial recruitment and four more at a later stage of the project.[4] Respondents who were interested in participating in the study were sent a consent form via email to sign and return by mail or fax. At this time, respondents were also asked to recommend articles related to Humanities computing they might like to see in the test journal. A shortlist was created from these suggestions and PDF documents of each article were found or created. These PDFs were then placed in the OJS Humanities computing Test Journal that had been installed for research purposes on the local network using a PostGreSQL database. Within the OJS test journal environment, the investigators were assigned roles as editors or managers. Each participant was manually enrolled in the system and sent email notification of his or her user name and password (which could be changed by the participant). This username and password allowed participants to use any of the reading tools that involved email, such as Send Link to Colleague. The test journal was then placed under a second level of password protection that would prevent the public from being able to access the copyright protected documents. 13

With a few exceptions, the interviews were conducted in the participants' private offices on their personal computers. In cases where the participants did not have a private office, interviews were held in a quiet public computer lab. During the interviews, the participants were asked whether or not they had been able to spend some time reading an article and exploring the reading tools before the interview. After returning to the article they read on the OJS, participants were taken through the reading tools one by one and asked a question, or series of questions, regarding the tool's value in enhancing their understanding of the article, evaluating the quality of an article, and providing materials for use in teaching or research. The interviews were summarized and the response to each tool categorized as either positive or negative. These responses were placed in a matrix[5] with, where possible, supporting extracts from the associated interview. 14

Results

We have organized the results of the interviews around a number of themes that emerged in the course of this study. These themes speak both to participants' reading processes and to the potential value of the tools. 15

Domain-Expert Readers Read for Context

Most participants felt that the About the Author and Author's Work tools helped them to evaluate the article they read by providing them with a basis for judging the position and contribution of the author. While a number of participants did recognize the author of the article they were reading, ten participants commented that if they did not recognize an author's name, they would be interested in gaining more information about the author's background and current research activities. Ten participants commented specifically on the usefulness of the About the Author tool in evaluating the article, and author, in question. Seven users also commented favorably on the Author's Work tool. While these domain-expert readers clearly trust their own judgment of an article, they are also interested in checking those judgments against easily ascertained signs or markers of expertise in the author's record. When using About the Author, a participant pointed out, "it can lend credibility to the article knowing who the author is, because – and for me specially this is an area that I don't know – I know the big names but I didn't know for example [this author's] name before reading this article" (P5). In response to the Author's Work tool, this participant stressed the importance of the venue in which the work was published but this was not the participant's only concern: "It's a question of credibility and knowing what he's working on. And the fact that he's published in other respectable venues[;] that he's somebody that does interesting stuff. I mean just reading a rather long abstract told me that . . . this is a guy who is doing important work" (P5). This telling comment reflects how a domain-expert reader arrives at a judgment by working through multiple signals, from the name-recognition of the journal to the careful assessment of an abstract from one of the author's other pieces. Two other participants P3 and P9 commented that they were more likely to look to the abstract than the metadata for certain kinds of information, such as keywords. It is worth noting that the abstract is a rather recent phenomenon in humanities journals (given this field's reliance on the literary tradition of the essay rather than the writing up of an experiment). This participant's ability to conduct a relatively quick check on the overall quality of the author's work, with the help of "a rather long abstract," speaks to the potential value of the abstract to the creation of indexing information when the full text is not available. 16

Paying attention to the "respectable venues" in which authors publish proved a common theme across a number of tools that involve searching the journal literature, including the Related Studies, Databases, and other tools. Eight participants commented on publication venue as being a factor in their evaluation of the author's article and other works. It became very apparent that where an article is published matters a good deal to these readers. Initially, a journal's peer-review process provided domain-expert readers with a pre-review and preliminary screening. And this pre-review – that is, prior to readers coming to their own judgment about an article – becomes all the more reliable through the readers' rating and rankings of journals within their field.[6] These readers place a great deal of emphasis on the journal's assessment of an article when performing their own evaluation: "I evaluate quality and significance partly by where things appear, which is the major problem that online publishing has" (P1). 17

The growth of online publishing has challenged many of the assumptions readers hold about "where things appear." Peer-reviewed work is appearing in new sort of a journal (e.g., *Borderlands eJournal: New Spaces in the Humanities*; *Computing in the Humanities Working Papers*). This new realm may pose a challenge to domain-expert readers who have been trained within a print culture: "Personally, I have [a] fair amount of skepticism of what's available there on the web unless I'm quite familiar with the sites and if they're discipline specific and created by academic institution or trusted institutions" (P12). There's reason enough for such skepticism. This wariness only adds to the work of arriving at a judgment about the article being read, as well as to the chance that one is wasting one's time with it. 18

The problem of online credibility has been compounded, particularly in the case of the humanities, by the relative slowness with which the standard journals in this field have become available on the Web. Many in the humanities today continue to believe that "online publication is still having real trouble getting accepted as... first rank" (P1). Fortunately, journal-hosting services such as Project Muse are now providing online access to a greater proportion of the peer-reviewed literature in the humanities even as the journals also remain in print for the present time.[7] Eventually, scholars will be able to settle back into their trusted ranking schemes. In the meantime, it is clear that the reading tools need to do more to identify the standards that govern its sources, as noted by one of the participants: 19

> What is peer-reviewed or not, we're not always sure. What is a good article or not in the field, it's very hard to define – much of it is . . . either password protected [or requires] pay. So, it's difficult to find and get to certain resources that might help me evaluate this article. Certainly, you can get to tools that might help to analyze this article, but evaluate – that's really a knowledge issue, not much of a tools issue. (P12)

The participant's point is well taken. The actual evaluation of the article depends not on the tools but on the reader's knowledge and experience. This point was reinforced by yet another participant in the study, who concluded that using the reading tools "reinforced an opinion [he] already [has] that throwing automated processes at text is not very valuable. Whereas having humans read text and make judgments and decisions about them is a) useful [and] b) what we're supposed to be doing" (P9). However, it might be noted that as the reading tools provide readers with related open-access peer-reviewed articles, the tools augment rather than hinder the reader's ability to "[make] judgments and decisions" about the work being read. Of course, more can be done to enable readers to gather background information readily on the sources used in the reading tools. As one participant pointed out, "that's what people check, they check editorial boards to see if they recognize the names and if they seem to be people from well known places" (P1).

The reading tools link to information about each database in order to provide readers with the means to evaluate each source. This link takes the reader to the database's homepage where in many, but not all, cases readers can discover more about how materials have been selected to be included in the database. While the About link works to describe the database, information about individual journals must come from other sources. Many online journals, however, carry with them the authority attached to their print-based origins. 20

Domain-Expert Readers Discuss Their Reading

For all their apprehensions about the status of online publishing, readers expressed support for other web-based means of communication such as online forums where scholars can gather for informal exchanges and commentaries. In the humanities, online discussions through such services as H-Net have long been a vital part of the field's use of online resources. The Online Forums tool directed readers to discussions thematically related to the article they read. Participants considered online forums to be more current and less "stilted" (P4) than other publication venues. Regardless of the article's publication date, the reading tools allow readers to tap into the most current discussions of the article's themes as well as any archived conversations. Online discussions reassured readers that they are part of a community. One participant commented, "you're not alone in front of your computer and that's why I like all these sites where you can comment and you can add and so on . . . it's very good to be part of a community on the internet" (P6). Another participant noted that this tool provided access to a level of engaged and incisive discourse that could not be found in the journals and described the value of such timely and current discussion in the evaluative process: 21

> There's open discussion and informal discussion on a wide variety of areas that you just won't find in some of the other academic publications . . .. It'll give you some insight on academic questions around the discipline that again you may not find in formal academic publications. So I think the Forums give you, in some ways, a more cutting edge insight into the people involved in the discipline, the things that are occupying them right now versus their book that published two years ago, three years ago, and who is within the community. (P12)

This participant made a connection between comprehending an article and understanding a field. The reader commented that seeing "other works by the author," as well as coming across "other people and disciplines who might be discussing works of this kind" and "discussions or arguments for this particular topic" in online discussion forums would "help you to understand your article" before concluding that "the more that you can find within this context . . . the more you will learn about the discipline and the more you are able to fit this article within that context" (P12). The community nature of an online forum could help users understand both articles and the field by fostering a dialogue with other researchers in the field, as they might take a moment to add to the discussion as a result of their current reading.[8]

Despite the favourable comments on online discussions, opinions on the Comment tool, which enables readers to leave comments on the article they were reading, were mixed. Five participants felt that being able to read opinions from other readers would be beneficial. One participant noted that leaving a comment could "create a scholarly conversation" (P5). Six participants, however, were skeptical as to how often they would use such a feature. As one participant noted, "I would very rarely want to bother making comments unless there's an active discussion already going on or there is likely to be an active discussion going on. I think the idea of accumulating layer after layer [of] interpretive, reactive stuff around articles is a bit, I mean it sounds cute, but how much of that stuff would anybody ever wade through in real life?" (P9). In a similar vein, while most participants seemed willing to email the author of an article with ideas they had about the work, the built-in Email Author tool was not felt to be necessary. One user, for example, commented that he would expect to find that link at the end of the author's biography rather than as a separate tool P12. Another participant suggested the possible use of an "add comment" feature that is configured to notify the author through email P4, revealing a familiarity with content management software that is increasingly common. 22

Domain-Expert Readers Are Teachers

A number of participants commented on the value of the reading tools for teachers and students. One participant stated, "I can see that some of these tools might be more useful to a student, than they would be to somebody who already has a fair understanding of the field" (P3). Another participant noted, "if I'm using [an article] for teaching . . . I like to put it in a larger context. So then [the Literary Critics tool] would help me relate it to other materials and to show how it was evaluated by other people" (P10). Other participants praised the Literary Critics tools as a valuable resource for students looking for supplementary material. One commented, "it's often very difficult to find good secondary references especially in a field like Humanities computing that's growing and in essence the last body of work [made digital] is criticism" (P12), while another felt that "being able to click on and find the other works of an author, the critic, that they're interested in is extremely helpful [for students]. And it helps [students] realize the scholarly context in which the person is working" (P5). Finally, a participant felt that access to dissertations would be of value to students who could use them to find out what research is currently being done in an area without waiting until the work reaches a published state P2. 23

Students were also seen as the major beneficiaries of the Define Terms and Dictionaries tools, which allow readers to look up individual terms in online sources: "I loved the [Define Terms] tool for student reading of academic writing and I would find that a wonderful tool, and I think I might really enjoy it for myself too in reading a more technically complex article" (P3). There was also a relatively rare instance of a participant seeing a crossover value for the tools between domain-expert and non-expert readers. For example, a second participant made it clear that she had already worked out a way to define terms readily online, but that these reading tools would benefit students P6. 24

The question of scholarly quality and restricted access was raised again, pointing to expert readers' emphasis on open access to research tools: "All you can do [with the Define Terms tool] is search Miriam [sic] Webster which is not a particularly 25

good dictionary...if I were looking up a word, I would go to the *OED* [*Oxford English Dictionary*] online which my university has access to" (P1). The same participant expressed similar concerns about the Dictionaries tool (which should rightly be renamed the Encyclopedia tool as was pointed out by P10), making it clear that the range of reference material made possible by library subscriptions was far superior to what was currently available through the reading tools. This shortcoming in the reading tool's ability to take full advantage of these readers' online resources is now in the process of being addressed.

Domain-Expert Readers Have Reading Methods

As domain-expert readers, participants exhibited techniques for augmenting and supporting the judgments they make about what they are reading. Some were already in the habit of turning to Google and Google Scholar. One asked, "rather than bothering going through all the separate tools [if] I wanted to find out books on the same topic, why wouldn't I just go to Google Scholar and be done with it?" (P7). A second participant put it plainly: "[rather than use the Author's Works tool] I would have done actually something simpler than going through all these searches . . . I would have gone through to Google Scholar or the online catalogue or library" (P6). Finally, a third participant commented, "so it's just giving me a different way of searching" (P2). The value of the tools for the domain-expert reader would seem particularly limited for those possessing more than a modicum of web savvy. Participants with less web experience were correspondingly less likely to compare the reading tools to Google. For them, however, the tools posed their own technical challenges. One participant commented that despite wanting "to be really positive here," we had to recognize that introducing these tools was "just making [his] life more difficult" (P3). 26

However, the participant who felt the tools just provided a different way of searching, which she knew how to do well enough, went on to pose something of a challenge for the design and reach of the tools. She pointed out that the tools did not open up new ways of reading for her; they were "a useful idea, but it doesn't change my perception of what's possible" (P2). This prospect of changing the domain-expert readers' sense of connections and contexts had been a hope with the tools. The change was intended in particular through the tools that connected the research article to government policies, media stories and instruction sites, and that connect readers to these themes in a focused way that is often difficult to obtain from broad Google searches. While none of these participants identified the Media Stories and Government Policy tools as particularly useful to their own research, which is understandable in most cases – "there's nothing in the government of Canada archives that's relevant to me on… Chaucer's text, right?" (P1) – four participants saw these tools as being of potential use to researchers in other fields. At least one participant recognized the value of these tools for the very difference it introduced: "I think that the media report, or the notion of having media articles there to see what the kind of public or the press is saying about the area is useful because it gives a very different perspective" (P2). 27

Domain-Expert Readers Generate Search Terms

Nine of the participants commented on the inadequacy of the provided search terms. These terms, you may recall, are normally drawn from the first two keywords the author provides for indexing the work in the process of submitting it to the journal, but in this case they were based on our reading of the article. Participants felt the terms provided were either too general to be useful or were actually misleading. In addition, one of the distinguishing features of domain-expert readers reading in their own field is how they "automatically constructed the main idea significantly more often when reading texts about familiar topics," which is to say that they are capable of generating very accurate keyword searches on their own [Afflerbach 1990, 43]. 28

One participant commented on the lack of focus in the search terms: "The key words both in the search scenario and in the metadata material I found almost too broad to be of assistance with this particular article" (P2). Another observed that viewing the metadata provided for her article "[explained] why most of the search tools later on failed as [abysmally] as they did, because it's improperly categorized" (P1). This participant added, "the direct link [to Google from the OJS] isn't so useful because the search terms in his case probably wouldn't be the ones I'd choose to use anyways" (P1). What needs to be noted here is that readers, before clicking "search" with any of the reading tools, are able to edit, delete or add search terms. Apart from overlooking this feature of the tools, one participant did go on to suggest that we should incorporate metadata automatically generated from the words used in the article P4.[9] Another felt it was the responsibility of professional cataloguers to read each article and determine categories and key words P5, which would greatly increase the accuracy of the reading tools, without question, while adding to the expense of running a journal. 29

Certainly, there is more that can be done to improve the indexing of the articles, and thus the precision of the searches. The journals' editors are able to provide authors with examples of indexing terms specific to their field, and they have the option of including a link to a thesaurus (e.g., ERIC Thesaurus in education) or classification scheme (e.g., Mathematics Subject Classification) that can further guide authors. It is in the authors' self-interest to learn more about how to accurately index their work. Another approach would be to draw on "folksonomy" tagging – used with blogs and social networking websites – in which readers can collaboratively suggest indexing terms for a work. These indexing terms can be used to create a word cloud which reflects the frequency with which each tag is used. Readers using the tools would then select tags or indexing terms from the cloud as their search terms. 30

Domain-Expert Readers Check References

One feature of reading research articles online that the project did not yet have in place when this study was conducted was found to be basic to the work done by researchers reading journal articles. Researchers check the works which the author references not only to see if the reference has been quoted and otherwise treated fairly, but because the references to an article provide one of the best sources of additional information on the topic. The reading tools operate independently of the article's references in the sense that the reading tools find related materials that have appeared after the article was published and that are potentially from a greater variety of perspectives on a topic than the list of references that an author assembles in writing an article. 31

Nonetheless, one participant in this study pointed out the value of having similar sorts of links from the reference in the article to the work referenced in the context of teaching a class on the original article: "[If] you have one text and the author is comparing it to another . . . you might zip over to have a look at the other text to see whether you want to discuss that one as well" (P14).[10] 32

Domain-Expert Readers Access Research Libraries

What became very clear from our interviews with the participants is that offering them the best of what was available in an open access format for their field fell well short of what they were used to, even in online environments. As faculty members and graduate students at research universities, they were able to log into their library's online resources from their offices and their homes, not to mention hotel rooms and airports. As a result, some participants compared the databases offered by the reading tools to those available through their library's subscription: "There are much better tools than the ones that are included [in the reading tools]. Now, if this included the MLA, then sure it would be useful, it's just that right now, the tools that are there aren't the ones that would give a complete listing of his work" (P1). One participant summed up the reading tools' open access limits this way: "I think it's an insurmountable problem . . . until this stuff really becomes free, because everything that's free is incomplete . . . and I care about complete" (P1). Dissertation Abstracts International, *Oxford English Dictionary*, and Project Muse were among the other subscription services that the participants felt were sorely lacking from the reading tools, given their dependence on open access databases. Another participant questioned the choice of the Internet Public Library as the only resource accessed by the Literary Critics tool and wondered why journal databases were not included P2. 33

However, there is a solution to this problem. To address this obvious and substantial shortcoming for domain-expert readers, the project is developing an OpenURL framework for the tools. This is what Google Scholar has done, enabling faculty members and students to register their library with Google Scholar, which then, with each search, identifies the items found that are in the library. It will take an extra step, but then so does logging into one's library from a remote location. The result will be, to use the words of one participant, a "comprehensive list [with] an indication of which items I can get at my desktop, and which ones I actually have to go to the library for" (P1). 34

Domain-Expert Readers Care About Access

At least a few of the participants showed great sensitivity to the access question. They were interested in providing a way for more than just scholars and students to be able to take advantage of these scholarly resources. That is, they were supporters of the *open access* movement in scholarly publishing: "What I'd hope [is] that . . . you could get academics hooked into the system so they'd publish copyright free articles under this system[;] in that [sense,] information will be free" (P4). Another participant drew the distinction of the university's obligation to the larger public, within the economic reality of publishing: "I'm very sensitive to the notion that publishing, even electronically, can cost money… but we're also working in a university environment where what we're doing is usually made available to the public; . . . the fact that I couldn't get this for free anywhere at all is problematic to me, philosophically speaking" (P5). This participant also stated on using this tool that "I think most of us . . . have a preference for . . . stuff that's open…. I don't object to having it on here and if people do want to pay I guess they ought to be able to [but] at the same time I would hate to encourage this sort of thing" (P5). There were also participants who responded negatively to the Pay-Per-View tool, which was the name of the only non-open access tool included among the reading tools: "I'd probably have to be pretty desperate to use [the Pay-Per-View tool] . . . even if we click on it, we don't seem to be getting very much" (P3). 35

Domain-Expert Readers Organize their Readings

As this was the participants' first experience with the reading tools and there were limits to how thoroughly the use of the tools could be reasonably explained to them, there were times when what the tools could do were misunderstood. Still, these misunderstandings provide insights into domain-experts' reading habits: 36

> I have my own way of organizing my information on my screen and one of the things would be to bookmark something. . . . so if I lose my ability to bookmark, because I'm within this window that's not giving me those features, then I can't save for myself a lot of material that I would want to be able to come back to. (P2)

The ability to bookmark the databases provided by the reading tools, as well as the individual articles found within any one database, could have been done at any point by using the browser's Bookmark feature, which of course operates independently of the reading tools. Still, this participant's interest in archiving and organizing the results of reading-tool searches does suggest that more could be done to facilitate this process. In an earlier version of the reading tools, readers could establish an online archive in which an article's metadata could be stored, whether readers were on their own computers or not. Highwire Press includes a Personal Archive link for articles as part of, for example, the *New England Journal of Medicine*. The point is that domain-expert readers are always in the process of building resources as a result of their reading. These readers keep an eye on the long-term utilization of the knowledge they encounter.

In Sum, Domain-Expert Readers Map Meaning and Value

What this study makes clear in relation to earlier studies of the cognitive processes that domain-expert readers apply to reading within their domain is that this reading process includes positioning and assessing the author and his or her ideas in relation to the larger field. The reader follows the flow of ideas within the article, all the while mapping and judging those ideas in relation to work with which they are already familiar and new work they have encountered through the reading tools. The judgment side of this process relies on finely tuned measures of hierarchical ranking. These measures are continuously recalibrated as the reader evaluates the author's other work, where the author's work is published, the author's institutional affiliation, and other external factors. This aspect of the reading process is aided by the use of the reading tools, much as the field-mapping and evaluative footnote championed by Edward Gibbon aided readers during the Enlightenment.[11] 37

Within the scope of this domain-expert reading process, a number of tools, such as Author's Work, Author Bio, and Related Studies, have proven capable of providing additional, readily accessible indicators or signals of the potential value of article they are reading. If not in every case, some readers used the tools to consult a number of sources, virtually on the fly, without losing sight, as it were, of the original article they were reading. Once having read the article, they may have come to reassess the initial judgments made on the basis of what they found through Author Bio and Author's Work. However, constant self-reflection is what distinguishes such readers [Wyatt et al. 1993]. Still, we should not assume that acts of reading scholarly articles always coincide with making sense of disciplinary fields, even for domain-experts. One participant drew this very distinction between comprehending the article and learning more about the field: "I read an article and I understand it . . . finding out more doesn't change my understanding of the article. It may change my understanding of the field, which is a bit different" (P1). The participant's reflections reiterate the interest of such readers in learning about who is doing what and where. This added contextual material makes for a complex reading, learning, and evaluation process. 38

For the domain-expert reader, learning to read a field means attaching ideas to people and vice versa. It means doing so within an evaluative framework that critically identifies not only who is associated with whom, but also which ideas are most promising and which are passé. These readers fix ideas to people, as well as to particular works, in part because this is necessary for utilizing the ideas within their writing and teaching; they are distinguished as readers, in other words, by their adherence to rules of attribution and to notions of intellectual propriety and property. Further design work on the reading tools, as has been suggested above, will seek to improve the tools' ability to assist domain-expert readers in learning more about the author and the work being read. The tools seek to augment, however slightly, these readers' deep knowledge of how their field is structured. In this way, the tools are intended to extend these readers' critical engagement with their field of expertise.[12] 39

Conclusions

We believe that with this new publishing medium comes an opportunity and an obligation. That is, it falls to researchers of the digital transformation of text to experiment with how online reading environments can extend more readers' critical engagement with text. The focus here, we would reiterate, is not on all forms of reading or all types of texts. The reading tools described here are not intended, for example, to diminish or displace what is achieved through *close reading*, as a major contribution of the humanities to scholarship. The reading environment at issue in this study, and in our related work, is that of the scholarly research article published in the online edition of a journal. What is clear is that journals have moved online. Although access is still limited, these journal articles are now open to an increasing number of readers. By providing a greater context for the reading of a given article, we can support the critical engagement of domain-experts, as well as increase the comprehension and interest of those reading outside their areas of expertise. 40

One participant in this study stressed the importance of reaching out to those working in the humanities in rethinking how we will read online, because of their very expertise in this area: "The people who need to work at this, need to be people who are very familiar with the book and its possibilities. . . . And a lot of those people . . . [are] staying away. Because they are familiar with the book and its possibilities and they don't really want to look at this" (P2). At the same time, there was at least 41

one participant in the study who was also imagining the possibilities for an improved reading environment. It took the form, for him, of a "highly developed information management forum that would incorporate a lot of these tools, but would really be towards . . . a kind of dialogue, a collective development of knowledge" (P4). Reading often leads to the formation of communities (which is one way of very briefly encapsulating the history of the university). Changes in reading habits and practices, on this scale, are not going to proceed in a simple or straightforward way; they are also not going to be dictated by website designs or research results.

Appendix 1

This appendix contains the full text of a report on the results of a study funded by the Canadian Social Sciences and Humanities Research Council, entitled *A Study of Professional Reading Tools for Computing Humanists*. 42

Executive Summary: A good deal of the emerging research literature concerned with online information resources focuses on information retrieval, which is concerned with the use of search engines to locate desired information. Far less attention has been paid to how the found materials are read and how that critical engagement can be enhanced in online reading environments. This paper reports on a study examining the question of whether a set of well- designed reading tools can assist humanities computing scholars in comprehending, evaluating and utilizing the research literature in their area. Thirteen computing humanists were interviewed regarding their experience using the reading tools. They were asked which tools, if any, and to what degree, these tools contribute to their comprehension, evaluation, and interest in utilizing the work they are reading. Responses varied widely among users but it was found that overall, the reading tools had the potential to lead to a variety of useful additional materials that would help one come to a better understanding of a particular article. The reading tools were deemed to be an exceptionally good resource for students or beginners in the field. Participants also identified several issues with the tools themselves and the web as a whole that affect the online reading and research experience. 43

Download PDF (864 pages). 44

Appendix 2

This appendix contains the original interview data on which this article is based. The source files are provided both as a compressed archive containing the original NVivo files, and also as plain text. 45

Download the interviews: as plain text, as a ZIP archive, as a tar archive. 46

Notes

[1]Additional studies are being conducted within the Public Knowledge Project on the value of such tools for readers who are not domain experts: e.g., Twomey, Willinsky and Quint-Rapoport, and Willinsky.

[2]The tools and databases in this second set were initially chosen and have been maintained by Chia-Ning Chiang, a doctoral student in information science, as part of the Public Knowledge Project team.

[3]The Humanities pose something of a challenge in terms of comprehensive searching, as the field lacks a "PubHum," comparable to PubMed, as a single comprehensive database for searching the life sciences.

[4]One interview was eliminated because of very poor sound quality in the interview recording and because the interview was incomplete.d

[5]The matrix is available with this article.

[6]Those working in the humanities are less likely than those in the sciences to track the ISI Web of Science citation index, which ranks journals within fields by their Impact Factor, based on the average number of times an article is cited across two years of a journal's publication. Many journals in the sciences advertise their impact factor, with authors closely attending to this ranking process, in the belief that it will affect the take-up of their work.

[7]Project Muse is a division of the Johns Hopkins University Press in collaboration with The Milton S. Eisenhower Library and currently offers over "300 high quality humanities, arts, and social sciences journals from 60 scholarly publishers," according to its website, http://muse.jhu.edu/.

[8]Not everyone participating in the study had an accurate notion of how online forums were being used, as one participant suggested that particle physics researchers "don't publish articles, they just go [to] online forums" (P4), thinking in all likelihood of http://arxiv.org, where particle physicists post their working papers, pre-prints, and post-prints, with such a level of daily activity among them on the site, including the posting of comments, that it must seem like an extended forum.

[9]This comparison of language is a feature that Google Scholar uses to find "related articles," along with co-citation and other comparative aspects, with an ability to use Google Scholar's "related articles" to be incorporated into a later version of the reading tools.

[10]In the next release of the journal software, Open Journal Systems, journals will be able to link an author's references to the article the author has cited through a number of strategies. The simplest of these is, in fact, a reading tool that takes the title of the cited article and feeds it into Google Scholar, a strategy that holds out the possibility for readers who do not have library access that there may be an open access copy of the article available online.

[11]Consider Keith Windschuttle's celebration of a new Penguin 3-volume edition of Gibbon's *History of the Decline and Fall of the Roman Empire* for, among other things, including all of the original footnotes: "Gibbon uses his footnotes not only to source his references but also to make lengthy, sometimes acrimonious, sometimes witty, commentaries on the veracity of both his primary and secondary sources. (A sample: 'The Dissertation of M. Biet seems to have been justly preferred to the discourse of his more celebrated competitor, the Abbé le Boeuf, an antiquarian, whose name was happily expressive of his talents.')."

[12]While this evaluative cataloging of people and ideas is not something that applies to those reading outside of their domain-expertise, we have found in other studies that a knowledge of the author also plays its part. For example, the author's background played an important role in the reading of non-expert readers, in terms of the prestige of the institution they work for, we found with policymakers [Willinsky 2003], while the name recognition from appearances at conferences came up in our research with Registered Massage Therapists and chiropractors [Willinsky & Quint-Rapoport 2007].

Works Cited

Afflerbach 1990 Afflerbach, Peter P. "The Influence of Prior Knowledge on Expert Readers' Main Idea Construction Strategies." *Reading Research Quarterly* 25.1 (1990): 31-46.

Alexander 2003 Alexander, P. A. *Expertise and Academic Development: A New Perspective on a Classic Theme*. Invited keynote address to the Biennial meeting of the European Association for the Research on Learning and Instruction (EARLI). Padova, Italy. August 2003.

Chen, Jing-Ping & Macredie 2006 Chen, S., F. Jing-Ping, and R. Macredie. "Navigation in Hypermedia Learning Systems: Experts vs. Novices." *Computers in Human Behavior* 22.2 (2006): 251-266.

Kopak 2006 Kopak, Rick, and Chiang Chiang. *Linking and Thinking: The Use of Typed Informational Relationships in Aid of Semantic Navigation*. Paper presented at ASIS&T Annual Meeting, Austin, November 2006.

Pressley 1995 Pressley, M., and P. Afflerbach. Verbal Protocols of Reading: The Nature of Constructively Responsive Reading. Hillsdale NJ: Erlbaum, 1995.

Rockwell 2011b Text Analysis Portal for Research(TAPoR). Geoffrey Rockwell, project leader. http://www.tapor.ca/.

Schreibman, Siemens & Unsworth 2004 Schreibman, Susan, Ray Siemens, and John M. Unsworth, eds. *A Companion to Digital Humanities*. Oxford, UK: Blackwell, 2004.

Siemens et al. 2010 Siemens, Ray, Analisa Blake, Cara Leitch, Karin Armstrong, John Willinsky, et al. *A Study of Professional Reading Tools for Humanists*. ETCL (Appendix 2 to this publication). 864 pp. Formerly at http://etcl-dev.uvic.ca/public/pkp_report/.

Stanitsek 2005 Stanitsek, Geog. "Texts and Paratexts in Media." Trans. Ellen Klien. *Critical Inquiry* 32.1 (2005): 27-42.

Twomey 2007 Twomey, Sarah. *Teacher Centered Professionalism: A Teacher Reading Group and New Forms of Access to Knowledge From a Feminist Perspective*. Unpublished dissertation, University of British Columbia, 2007.

Willinsky & Quint-Rapoport 2007 Willinsky, John and Mia Quint-Rapoport. "How Complementary and Alternative Medicine Practitioners Use PubMed" *Journal of Medical Internet Research*, 9.2 (2007) http://www.jmir.org/2007/2/e19.

Willinsky 2003 Willinsky, J. "Policymakers' Use of Online Academic Research." Education Policy Analysis Archives11.2, 2003, http://epaa.asu.edu/epaa/v11n2/.

Windschuttle 1997 Windschuttle, Keith. "Edward Gibbon and the Enlightenment." *The New Criterion* 15.10. June 1997. http://www.newcriterion.com/archive/15/jun97/gibbon.htm.

Wineburg 1991 Wineburg, S. "Historical Problem Solving: A Study of the Cognitive Processes Used in the Evaluation of Documentary and Pictorial Evidence." *Journal of Educational Psychology* 83.1 (1991): 73-87.

Wineburg 1998 Wineburg, S. "Reading Abraham Lincoln: An Expert/Expert Study in the Interpretation of Historical Texts." *Cognitive Science* 22.3 (1998): 319-346.

Wyatt et al. 1993 Wyatt, D., M. Pressley, P.B. El-Dinary, S. Stein, and R. Brown. Comprehension Strategies, Worth and Credibility Monitoring, and Evaluations: Cold and Hot Cognition When Experts Read Professional Articles That Are Important to Them. *Learning and Individual Differences* 5 (1993): 49–72.