JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008

Computer-Supported Collaborative Learning
DOI 10.1007/s11412-008-9040-x

# Annotations and the collaborative digital library: Effects of an aligned annotation interface on student argumentation and reading strategies

**Joanna Wolfe**

Received: 13 August 2007 / Accepted: 11 February 2008


**Abstract** Recent research on annotation interfaces provides provocative evidence that anchored, annotation-based discussion environments may lead to better conversations about a text. However, annotation interfaces raise complicated tradeoffs regarding screen real estate and positioning. It is argued that solving this screen real estate problem requires limiting the number of annotations displayed to users. In order to understand which annotations have the most learning value for students, this paper presents two complementary studies examining the effects of annotations on students performing a reading-to-write task. The first study used think-aloud protocols and a within-subjects methodology, finding that annotations appeared to provoke students to reflect more critically upon the primary text. This effect was particularly strong when students encountered pairs of annotations presenting different viewpoints on the same section of text. Student interviews suggested that annotations were most helpful when they caused the reader to consider and weigh conflicting viewpoints. The second study used a between-subjects methodology and a more naturalistic task to provide complementary evidence that annotations encourage more reflective responses to a text. This study found that students who received annotated materials both perceived themselves and were perceived by instructors as less reliant on unreflective summary strategies than students who received the same content but in a different format. These findings indicate that the learning value of an annotation lies in its ability to provoke students to consider and weigh new perspectives on the primary text. When selected effectively, annotations provide a critical scaffolding that can support students' critical thinking and argumentation activities. Collaborative digital libraries and applications for the Web 2.0 should be designed with this learning framework in mind.

**Keywords** Annotation · Anchored discussion · Digital libraries · Reading interfaces · Reading-to-write · Computer supported argumentation · Persistent conversation

J. Wolfe (✉)
University of Louisville, HM 315, Louisville, KY 40292, USA
e-mail: joanna.wolfe@louisville.edu

## Introduction

In the humanities, social sciences, and other disciplines, much learning is conducted by assigning primary texts—such as newspaper articles, historical documents, argumentative essays, or fiction—and discussing these texts as a class. Computer-mediated environments are often used to facilitate these discussions, with email, chat, and discussion boards the most common tools in use. Such discussions of primary texts are not limited to the academic realm either: most newspapers and magazines host online discussion boards where readers can comment on recent publications; likewise, many digital libraries are seeking to create online spaces where readers can discuss collection materials. These public spaces attempt to foster reading communities by providing public forums for democratic debate and discussion.

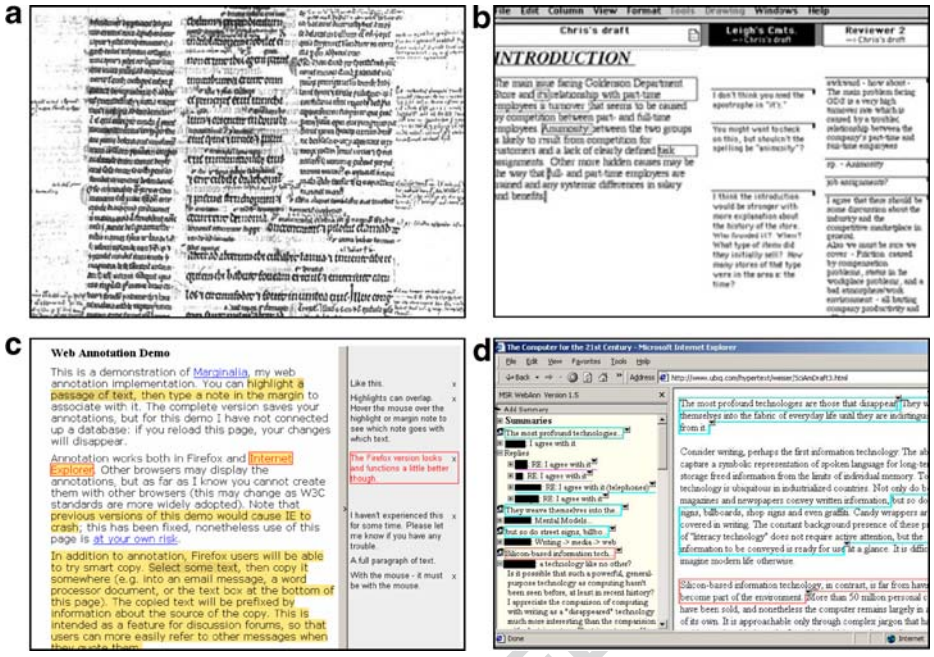
Unfortunately, as Kaplan and Chisk (2005) note, the chat or discussion board interfaces generally used for such discussions require a visual separation of primary text and commentary that can make it difficult to integrate the two. The cognitive effort required of readers who have to switch their attention back and forth between two separate visual panes (primary text and commentary) can be quite considerable. Moreover, in order to make a comment, users of discussion board interfaces must reconstruct the context they are responding to (Brush et al. 2002): users must not only identify the text being discussed, but may also need to identify the specific paragraph, line or word to which their message pertains. This difficulty of making document references in some online environments reduces the overall helpfulness and quality of reflection in these conversations (Honneycutt 2001).

To solve the cognitive burden produced by systems that place primary text and commentary in separate visual frames, a number of developers have turned to annotation interfaces to support discussions on primary text (Brush et al. 2002; Glass 2005; Kienle 2006; Kaplan and Chisk 2005; Lebow and Lick 2002; Ovsiannikov et al. 1999). An annotation interface is one in which readers comment in the margins or directly upon a primary text much as readers may write in the margins or between the lines of privately owned print texts. This "marginalia" is then shared with other readers of the primary text in order to support discussion. Annotation interfaces are sometimes referred to as anchored discussions (Brush et al. 2002; van der Pol et al. 2006) because comments are not only presented in the same visual pane as the primary text, but are also clearly linked—or "anchored"—to a specific passage in the primary document. Thus, annotation interfaces clearly help users establish a context for a comment. Figure 1 shows examples of four different systems that use annotation interfaces to support discussions of primary texts.

Research comparing annotation interfaces with discussion boards supports the theory that the ability to easily anchor annotations to primary text reduces the cognitive effort required to clarify the context of a comment and can improve the overall quality of conversation (Brush et al. 2002; Guzdial and Turns 2000; van der Pol et al. 2006). Brush et al. (2002) for instance found that when students used an annotation interface they made more comments on technical articles, were more likely to reply to other students' comments, and made more specific comments than when using the discussion board. Van der Pol et al. (2006) similarly found an increase in commentary when students used an annotation interface and moreover found that the annotation interface increased task-directedness, encouraged rereading of the primary text, and produced more meaning-oriented discussion. These researchers conclude that annotation systems produce more constructive collaboration centered on understanding the meaning of the primary text than

JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008






Q2

**Fig. 1** Four annotation systems. **a** Bible with commentary from the *Glossa Ordinaria*, ca. 850–1499. M389. Annotations are written in all *margins* and interspersed between *lines* of the primary text. Such annotations were copied along with the primary text and supported debates across centuries. **b** PREP editor uses an infinite expanding margins interface to support commentary by multiple authors. Although developed as a tool to support peer review on documents, such an interface could easily be used to support discussions of primary texts. **c** Marginalia allows readers to select text and enter a corresponding annotation in the right column. All comments appear in a single column, making it difficult to align text and comments. **d** WebAnn was designed to replace a discussion board. Users select text and enter comments in the left column where a threaded discussion interface allows other readers to respond



discussion forums, which produced a more debate-like, individual opinion-centered conversation.

In sum, research on annotation interfaces provides some compelling (though far from conclusive) evidence that anchored, annotation-based discussion environments may lead to better conversations about a text. Unfortunately, annotation interfaces can be very difficult to plan and design, especially when compared to discussion boards, and designers have a number of options to weigh and consider. The next section presents some of the basic design options, reviews research suggesting that aligned annotation interfaces (such as those in Fig. 1) may be preferable to other display methods, and sketches some of the complex design decisions that aligned annotation interfaces raise and which we have only just begun to understand.

## The design tradeoffs of annotation systems and the problem of screen real estate

One of the major issues in developing annotation systems involves deciding how to position the annotations vis-à-vis the primary text. This is a thorny issue not only because some positioning methods require complex programming algorithms to display correctly, but also because screen real estate is limited, creating problems when multiple annotations

are anchored to the same place in the primary text. Screen real estate becomes a particular problem when designers must decide how (or whether) to support annotations upon previous readers' comments.

Annotation systems have tended to use one of following basic formats for annotation display:

- Footnote or split screen. Annotations appear at the bottom of the screen or at the bottom of some unit of text, which may be the document, the page, or the paragraph. Thus, the annotation is clearly anchored to a specific section of primary text but does not appear in the same visual frame as the primary text. Examples include CaMILE (Guzdial and Turns 2000).
- Interlinear. Annotations are inserted directly in the primary text.
- Mark-up or "sticky note." Annotations are inserted in layers over the primary text. These layers often operate on a "sticky note" or "Post-It note" metaphor where the annotations appear as notes that can be moved around to different locations on the text. Examples include Aleph (Kaplan and Chisk 2005), iMarkup and Adobe Acrobat.
- Aligned or marginal. Annotations are placed in a column—or "margin"—near the primary text to which they correspond. A minimal amount of highlighting is generally used to help anchor the comment to the primary text. The four annotation systems in Fig. 1 all present different types of marginal alignment systems.

Research on annotation interfaces suggests that while the footnote and interlinear layouts are by far the easiest to program, they are also the least useful to readers. Zellweger et al. (2000) found that footnote comments placed at the bottom of the screen were frequently missed by users, a result that is lent support by the finding of Wojahn et al. (1998) that a split-screen footnote interface reduced the amount of communication about problems in a text and the finding of van Ostendorp (1996) that a footnote interface resulted in less effective notetaking. These findings make sense theoretically since a footnote interface places a high overhead on users by requiring their eyes to travel long vertical distances across the screen in order to reconcile primary text and annotation.

Similarly, interlinear and mark-up, or "sticky note," interfaces—while potentially useful for commenting on documents-in-progress that an author is going to revise—do not seem to be useful layouts when the purpose is to read and discuss a published, primary text. In their study of annotation layouts, Zellweger et al. (2000) found that many readers strongly objected to annotations that interrupted the flow of the primary text. Cabanac et al. (2007) moreover note that interlinear annotations have the potential to be confused with the primary text. As the number of annotations on any given text grows, one would imagine that such interruptions would become increasingly annoying: thus, readers who might tolerate the occasional annotation interrupting the primary text might quickly become intolerant when the quantity of annotation begins to rival the amount of primary text, as could become the case with highly controversial or particularly salient texts.

Aligned interfaces (where annotations appear in the "margin" of the text) thus theoretically appear to offer the best design alternative for annotation interfaces because they allow close positioning between primary text and commentary without obscuring or interrupting the primary text. However, once designers start to implement this interface, problems arise: the margin after all is a very finite space and annotators often have a lot to say. How can designers maintain the alignment between annotation and primary text when there are large numbers of annotations to display? How should the screen adjust to display

JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008


commentary upon a previous readers' annotation? In short, how should the system respond to the considerable problem posed by screen real estate?

Limited screen real estate is perhaps the central obstacle in the development of aligned annotation interfaces, and designers have yet to propose a fully satisfactory solution to this problem. The four interfaces in Fig. 1 display four different approaches to the screen real estate problem:

- The medieval manuscript (Fig. 1a) allows readers to write comments in any of the four margins of the page as well as between the lines of the primary text, and the annotations are closely aligned with the primary text. While this interface makes good use of limited page real estate, the amount of annotation allowable is clearly limited by the size of the page. Medieval scholars resolved this problem through manual editorial control: when a manuscript was recopied the annotations may be edited or removed.
- PREP Editor (Fig. 1b) similarly positions commentary on the same line as the anchoring primary text and uses an "infinite expanding margin" interface that allows the commentary to expand beyond the confines of a single page (Wojahn et al. 1998). The comments of each individual reader/reviewer are displayed in a separate column. Users can hide and reposition the columns, and the system supports annotations upon annotations. Despite the apparent elegance of this system, the need to restrict each column to a specific individual makes this interface unwieldy once more than three individuals have commented.
- Marginalia (Fig. 1c) places all readers' comments in a single margin. However, once again the interface becomes cluttered once a large number of annotations are present; moreover, readers cannot directly comment on one another's comments (Glass 2005).
- WebAnn (Fig. 1d) also uses a single margin solution combined with a threaded discussion interface to support discussion of previous readers' comments (Brush et al. 2002). However, in order to solve the screen real-estate problem, the designers have resorted to only displaying the first line of each comment: readers must click in order to see the full text.

As more and more annotations are added, more cognitive effort is required from users to reconcile annotation with primary text. Thus, as the popularity of the annotation system increases, its primary advantage (allowing users to move effortlessly between primary text and annotation) decreases. It seems unlikely that a satisfactory layout solution to this real-estate problem will be found: either designers require users to take some action, such as clicking or scrolling, to see the entire annotation, or they present users with a cluttered interface that requires mental effort to "line up" primary text and annotation. Thus, some other solution to the real-estate problem is needed.

In order to continue to reap the benefits of an aligned annotation system as it grows in popularity, designers need some way to limit the number of annotations displayed at any given time. Such a solution requires understanding which annotations are the "best" and designing a system that prioritizes these best annotations by displaying them first. Preliminary work on developing algorithms to rank the quality of individual annotations has already been initiated by Cabanac et al. (2007). However, before we can really begin such work in seriousness, the design community needs a better understanding of what quality means in this context. In other words, what makes a given annotation "good"? And, more particularly for CSCL designers, what makes a given annotation good for student learners? How can we assess the learning value of a given annotation?

## Annotations and the collaborative digital library

This research takes the collaborative digital library as the context for understanding the question of what makes an annotation good for student learners. A considerable body of research in digital library communities has focused on how to capture the types of annotations that readers often make in the margins of books and other materials and repurpose them for other scholars and learners (c.f., Marshall 1997, 1998; Wolfe and Neuwirth 2001; Marshall and Brush 2004; Kaplan and Chisk 2005). As libraries are offering a growing number of their holdings in electronic format and as technologies for annotating these materials are improving (particularly with Tablet PCs), it is only a matter of time before this loop is closed and readers will be able to share these individual annotations with other library users. Even if only a small proportion of these annotations are eventually shared (Marshall and Brush 2004), their presence will dramatically increase the opportunities for collaborative learning in digital libraries.

It is worth noting, as Fig. 1a makes clear, that this practice of sharing marginalia is hardly new. Medieval readers routinely added commentary to the margins and interlinear spaces of these manuscripts. Because a single physical manuscript was extraordinarily expensive, it was shared with a large community of readers and any annotations made were automatically public. These annotations were places where debates about the primary text raged in persistent conversations that often spanned centuries. Moreover, annotations were also pedagogical in that they were crucial in teaching students how to read by inducting them into the thoughts and habits of a particular community of readers. In fact, as Jackson (2001) makes clear, the practice of circulating annotated books within specific communities of readers was common up until the large-scale mass production of books early in the twentieth century.

CSCL designers should keep these socializing and training purposes in mind when developing annotation systems. As annotations become increasingly important to digital libraries, they offer not only new ways to respond to materials, but also new ways to read and to have our readings shaped by the thoughts of others (Kaplan and Chisk 2005; Marshall 1997). However, most research has evaluated annotation systems based upon the comments that readers produced (Brush et al. 2002; Wojahn et al. 1998; van der Pol et al. 2006) rather than examining the effect that encountering others' annotations might have upon learners' reading practices or their perceptions of the primary text.

The research reported below assesses the learning value of annotations such as might be encountered in a collaborative digital library or the Web 2.0. In particular, these two studies attempt to address the following research questions:

1. How do annotations shape reading practices? What learning value do annotations offer readers?
2. How can we determine which annotations might be particularly beneficial for learners? How can developers rank or prioritize annotations?
3. What design and pedagogical implications can we infer based upon the answers to these questions?

Two studies with complementary methodologies are presented here. The first study uses think-aloud protocols and a within-subjects methodology to examine how annotations affect reading practices of both novice and advanced student readers. Because think-aloud protocols are a labor-intensive research method involving time-consuming data collection, transcription, and coding, a small number of participants are examined. This small participant pool is offset by the rich analytic detail this method allows. The second study

JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008


complements the first by involving a greater number of students in a much more naturalistic task—composing a written response to assigned readings as class homework. This study uses a between-subjects design to compare the essay quality and self-reported writing plans of students who received annotations with those who received similar content in a different format. Thus the two studies work together to answer the question of what annotations offer learners: the first study provides rich insights into how annotations affect students' reading and thought processes while the second study provides less detail but more methodological rigor.

## Prestudy: Annotating *The Computer Delusion*

To prepare for the main studies below, 12 college composition instructors were first asked to annotate an excerpt from "The Computer Delusion," a popular essay from *Atlantic Monthly* arguing against the emphasis on computers in K-12 education (Oppenheimer 1997/1999). This essay was chosen because it has appeared in at least one popular composition course anthology (Lunsford and Ruszkewicz 1999) and is therefore typical of material students are asked to read and write about in composition courses. Moreover, earlier pilot studies using this essay found that the issue is one that college students find interesting and controversial with approximately half of student participants agreeing with the essay's cautionary arguments and the other half disagreeing.

In addition to the essay, instructors received four letters to the editor of the *Atlantic Monthly* and a short profile piece emphasizing the benefits of technology in college classrooms appearing in *The Chronicle of Higher Education*. These auxiliary materials were included to present a range of viewpoints on the issue.

Instructors were asked to read and annotate the materials as if they were planning to write a persuasive argument on the role of computers in education. After annotating the materials, the instructors went back and identified their reasons for making the annotations. The 12 instructors made a total of 284 annotations containing at least one word. Of these, 22% were negative evaluations of the primary text, 27% were positive evaluations, and 35% represented attempts to comprehend the material.

Prior research suggested that annotations representing strong viewpoints influence students more than those with neutral comments (Wolfe and Neuwirth 2001). Therefore, from the instructor annotations, a subset of 15 representative "pro" and 15 representative "anti" annotations were culled. These 30 annotations were selected based upon their typicality (multiple instructors made similar types of comments on the text), their lack of idiosyncratic references intelligible only to the annotation author, and their succinct expression of a clear position or viewpoint on the primary text.

The 15 "pro" annotations all agreed with the basic premise of the essay and included comments such as

- "Yes. Great point."
- "All very persuasive quotations."
- "Academic source—very credible evidence."
- "Yes—there's too much emphasis on the cutting edge rather than fundamentals."

The 15 "anti" annotations clearly disagreed with the premise of the essay and included comments such as:

- "I disagree—the internet is global, not linear"
- "I find this hard to believe"

- "What evidence is there for this claim?"
- "But talking on the phone or writing a letter are also isolating—should we get rid of these too?"

These 30 annotations then formed the basis of the materials used in the studies. Study 2 used all 30 annotations while study 1 (with an abbreviated set of materials, necessary to avoid tiring the think-aloud participants) used 14 of these annotations. These annotations were printed in the margins of the primary text, next to their anchor, which was underlined so that readers could clearly match up primary text and associated commentary. The "pro" annotations were all associated with the initials "D.H." while the "anti" annotations were associated with the initials "R.W." Thus, the materials represented a context in which two individuals with very different perspectives had annotated the primary text. The annotations by the two hypothetical individuals were presented in different fonts and were positioned slightly differently in the marginal column. See Fig. 2 for examples of this formatting. An introduction to the study informed students that these two annotators were composition instructors who were planning on assigning the materials to their classes. This explanation of the annotators was intended to provide some credibility to the annotations.

## Study 1: Within subjects think-alouds

### Method

#### *Participants*

Seven students—two advanced and five novice readers—participated in this study. The five novices—two women and three men—were all first-year freshman students recruited at the end of the term from their freshman composition courses at a mid-western, urban university. The two advanced readers—both women—were recent graduates with degrees in English who had both been admitted to graduate school.



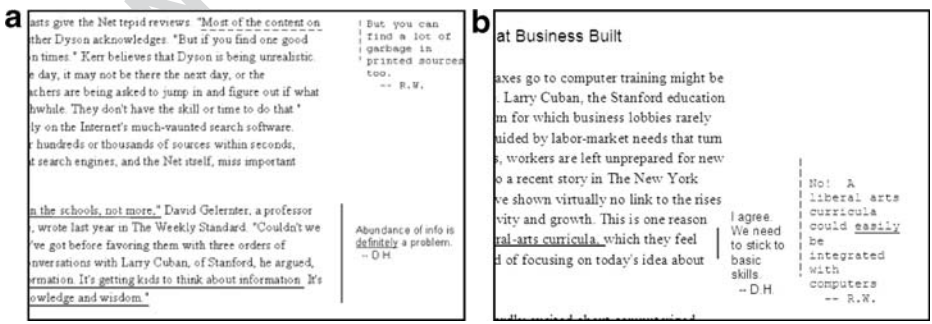



**Fig. 2** Aligned annotation condition of the materials. **a** Paragraph with a single con annotation followed by a paragraph with a single pro annotation. **b** Paragraph with both a pro and a con annotation. In all cases, annotations were anchored to the primary text with *underlining* and a *vertical line* in the margin. The comments of two annotators were differentiated from one another, as is common in annotation systems (see Figs. 1b,d)

JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008


*Materials*

Condensed versions of the *Atlantic Monthly* essay by Todd Oppenheimer described above were used for the think-aloud protocols. The condensed materials were necessary to avoid tiring the readers as they thought aloud. The materials were structured so that the essay alternated among paragraphs with no annotations, paragraphs with a single pro or con annotation, and paragraphs that contained both a pro and a con annotation.

In all, the essay contained 12 annotated paragraphs and 12 non-annotated paragraphs. The annotated paragraphs contained five paragraphs with a single "pro" annotation in support of the primary text, five paragraphs with a single "con" annotation that disagreed with the primary text, and two paragraphs that contained both a "pro" and a "con" annotation (see Fig. 2). Annotations were not labeled as "pro" or "con," however. Because of the way in which the data was collected, the two advanced readers received a slightly different, but equivalent, version of the text.

To ensure that the annotated and non-annotated primary content was similarly provocative, some passages that had been annotated by over 40% of the instructors were presented here without annotations. This guaranteed that some clearly provocative content was presented without annotation. Overall, annotated and non-annotated paragraphs were roughly equivalent in length with an average of 132 words per annotated paragraph vs. 117 words for non-annotated paragraphs.

*Think-aloud prompt and procedures*

Participants were first informed of the think-aloud procedure and then given a task assignment sheet that instructed them to read the materials and write "a response essay that takes a position on these materials." Participants were told to follow their normal reading and writing processes as much as possible with the exception that they were to read aloud and speak everything that passed through their minds. If participants fell silent, a researcher prompted them by asking "what are you thinking now?" Because the focus of the study was on reading practices, protocols were stopped after 45min—before some participants had time to complete the essay. This was done to avoid tiring participants. The think-aloud protocol was then followed by an interview that asked participants about their general reading practices and background and their specific strategies in approaching this assignment. The interviews ended by asking participants specifically about their thoughts on the annotations.

*Analysis of think-aloud protocols*

The comments participants made while reading *The Computer Delusion* were segmented into continuous episodes for analysis. An episode is defined as a unit of concentration in the reader's process: a new episode begins whenever the reader shifts focus or changes a train of thought (Charney 1993). These episodes were then categorized according to a coding scheme based upon a modified version of that developed by Charney (1993). The three basic types of comments were: comprehension, evaluation, and other. These three coarse-grained categories were then broken down into further fine-grained subcategories. Table 1 provides a detailed description of the coding scheme. Since only the fine-grained analysis of evaluation comments is reported here this is the only category described in detail. A second rater analyzed a random subset of 20% of the episodes. Inter-rater agreement on the three categories of comprehension, evaluation and other was $K = 0.87$ using Cohen's simple

**Table 1** Coding schema for think-aloud protocols

| Comment type | Description | Examples |
|---|---|---|
| Comprehension | Indicates the reader is working to understand the primary text. Includes summarizing, paraphrasing, rereading, drawing logical inferences from the text, and commenting on one's own understanding | He thinks that you have to develop knowledge of the hands before moving to computers |
| | | So they're saying that computers can't create new ideas |
| | | I don't know what that means |
| Evaluation | | |
| Credibility | Assesses the credibility of the text's sources or evidence | And she's pretty credible because she works for LucasArts, which makes a lot of games |
| Agree | Indicates agreement or support of a claim made by the primary text, a quotation, or an annotation. | I agree with that. |
| | | That's a good point. |
| Disagree | Challenges a claim or indicates disagreement with a claim made by the primary text, a quotation, or an annotation. | I think that's a bit ludicrous |
| | | That's not a very good argument |
| Preparation | Takes an explicit position that might appear in the writer's essay *and* does not simply echo what has been said in the main essay | I think teachers need more training to avoid these problems |
| | | Actually, I see a way that they could do both |
| Other | The reader's intention is unclear, the comment is focused on the reader rather than the task, or the comment is irrelevant | Hmmm |
| | | I'm thirsty |
| | | I'm just going to flip back here. |

kappa. Inter-rater agreement on the fine-grained categories of credibility, agreement, disagreement, and preparation was $K = 0.80$, also using Cohen's simple kappa.

## Results of think-aloud study

### *Quantity of comments*

The seven students made a total of 675 comments as they read the materials and thought aloud. The number of comments made by individual students ranged from a low of only 12 comments to a high of 161 comments. Overall, novice readers made more comments than advanced readers (see Table 2).

Table 2 shows that readers made more think-aloud comments on paragraphs with annotations than paragraphs without annotation—particularly when the annotation contained some type of critical commentary. Overall, annotated paragraphs provoked approximately 50% more comments than non-annotated paragraphs, a difference that was statistically significant, $F(1,167) = 8.79$, $p < 0.01$ (degrees of freedom reports the total number of paragraphs analyzed). Moreover, the type of annotation appeared to influence

JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008



**Table 2** Average comments per paragraph (and SD) for advanced and novice readers t2.1

| Condition | Novice comments per paragraph | Advanced readers' comments per paragraph | Marginal means | |
|---|---|---|---|---|
| | | | | t2.2 |
| No annotations (*n*=12) | 3.8 (3.2) | 2.7 (1.7) | 3.5 (2.9) | t2.3 |
| Annotations | | | | t2.4 |
| Pro annotation (*n*= 5) | 4.2 (3.2) | 3.2 (2.4) | 3.9 (3.0) | t2.5 |
| Con annotation (*n*=5) | 5.9 (5.2) | 4.4 (2.5) | 5.5 (4.6) | t2.6 |
| Both pro and con annotations (*n*=2) | 8.1 (5.3) | 5.0 (2.5) | 7.2 (4.8) | t2.7 |
| All annotated paragraphs (*n*=12) | 5.6 (4.6) | 4.0 (2.5) | 5.1 (4.2) | t2.8 |

readers, $F(3,167) = 6.12$, $p < 0.001$. Readers were significantly more likely to comment if the paragraph contained both a positive and negative annotation than if the paragraph contained only a positive annotation (Duncan's post hoc analysis, $p < 0.01$). These findings suggest that annotations increase the amount of reflection student readers engage in.

*Types of comments*

Not only did annotations appear to influence the quantity of comments, but they also influenced the type, or quality, of comments. Figure 3a and b show that the type of annotation influenced the number of evaluations both novices and advanced readers made, $F(3,167) = 6.46$, $p < 0.001$. This shift is particularly remarkable in the comments made by the novice readers. When no annotations were present, the novice readers focused primarily on understanding the text: 65% of their comments involved comprehending the text and only 30% evaluated the text by agreeing or disagreeing with its claims. However, with the two paragraphs containing *both* positive and negative comments (see last column of Fig. 3a) this pattern was reversed: the majority of the readers' efforts (55% of comments) consisted of evaluating the text and only 40% of their comments focused on comprehension.

This shift from a focus on comprehension to a more reflective focus on evaluating the text did not, however, mean that readers' were expending less effort working to understand the text. If anything readers exerted more effort verifying that they comprehended the text when negative annotations were present (i.e., in the *con* and *both* conditions), although this difference is not significant. These findings suggest that annotations can productively influence novice readers' active reading practices.

Paragraphs containing *pro* annotations only did not have nearly as strong an impact as paragraphs that contained some sort of negative annotation. In fact, although the data is far from significant, it does suggest the possibility that positive annotations could reduce the amount of effort readers put into comprehending the primary text. Figure 3a shows a slight dip in the number of comprehension comments per paragraph in the single *pro* annotation condition that may suggest that *pro* annotations encourage readers to assume that they understand the material. Additional research is needed to test this hypothesis.

Wes, one of the novice readers, provides an excellent example of how annotations influenced readers. Most of Wes's comments while reading the primary text of Oppenheimer's essay were concerned with vocabulary (i.e., whether or not he understood certain words). By contrast, after reading an annotation expressing strong support for Oppenheimer's position, Wes comments, "I would have to agree with that. Completely, yeah." At another point, when encountering two annotations expressing conflicting viewpoints, Wes responded by articulating his own stance on the materials. Wes paused

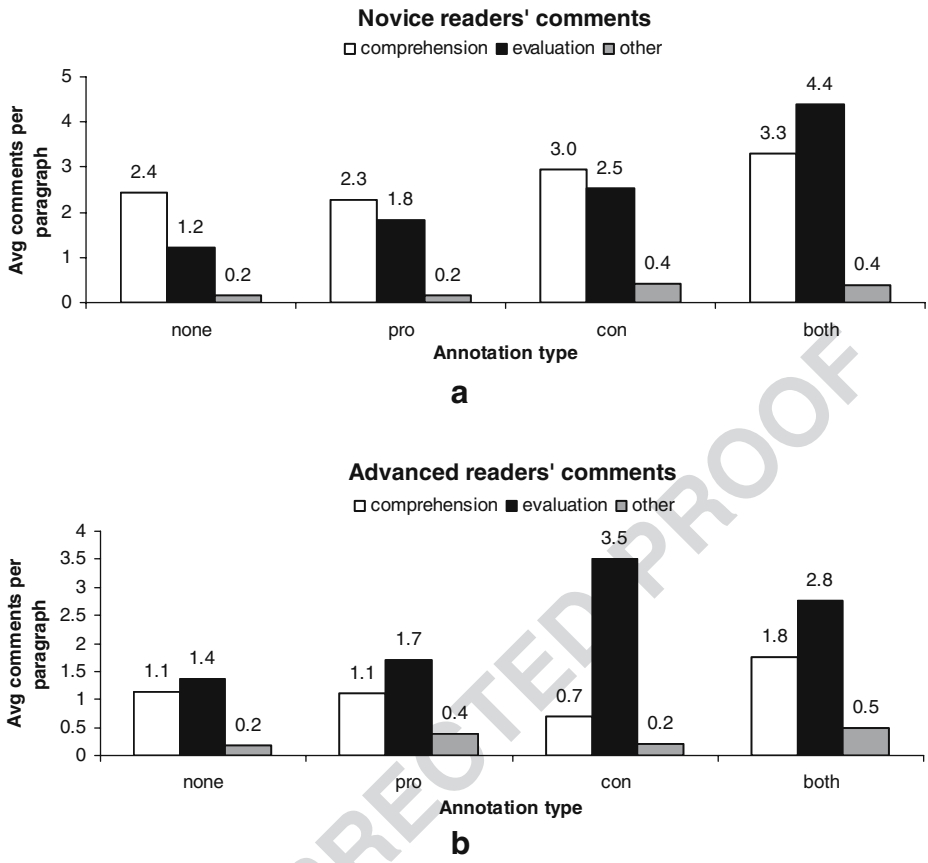



**Fig. 3** **a** Novice readers' mean number of comments per paragraph by type of annotation. **b** Advanced readers' mean number of comments per paragraph by type of annotation



only four times to make such evaluative comments—and three of these were in direct response to annotations.

Other readers seemed to respond to the annotations as invitations to work out conflicting points of view. For instance, Eli, another novice reader, spent nearly a minute after reading a strongly opposing annotation (*con*) reflecting on the costs and benefits of different communication technologies, ending this reflection by saying "so that's how you see it from the other point of view." Toby, another novice reader, responded to nearly every annotation by stating whether he agreed or disagreed with it—and such initial evaluations were often followed up with additional reflection on his own stance. Thus, the annotations appeared to provoke these novice readers to take a stance on the text.

*Fine-grained analysis of evaluations*

Not only did readers perform more evaluations on annotated paragraphs, but a more fine-grained analysis suggests that readers' evaluations became more complex in the *con* and *both pro* and *con* annotation conditions. Figure 4 shows that the type of annotation influenced the number of reader comments that agreed with a claim ($F(3,167) = 7.00$,

JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008



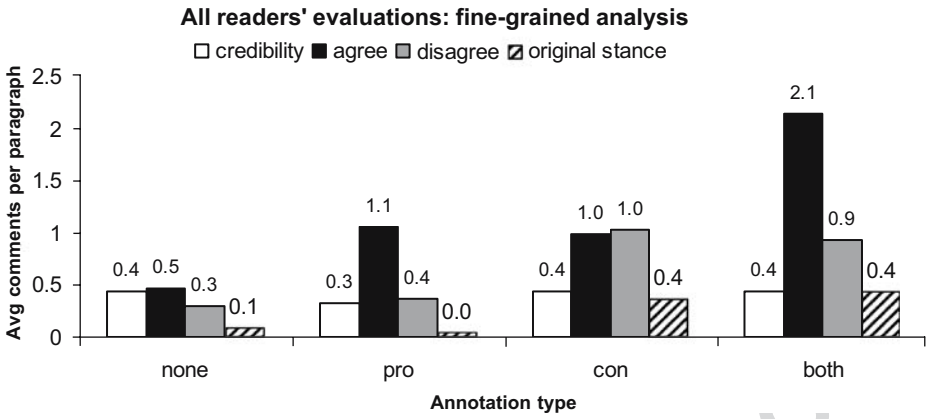



**Fig. 4** All readers' mean evaluation comments per paragraph by type of annotation

$p < 0.001$), the number of comments disagreeing with a claim ($F(3,167) = 6.49$, $p < 0.001$), and the number of comments that reflect new arguments that do not just simply echo what has been said in the main essay ($F(3,167) = 3.33$, $p < 0.05$). Post-hoc analyses suggest that the *both pro* and *con* annotation condition was significantly more likely than the *none* and *pro* annotation conditions to provoke readers to agree or disagree with the materials or to articulate an original argument on the topic. When two annotations—one positive and one negative—were appended to a paragraph, readers made over three times as many evaluations and articulated over five times as many original arguments as they did when no annotations were present. These findings therefore suggest that the annotations may have prompted the students to step back from the text and take a position on the argument.

Tabetha, one of the novice readers, exemplifies how paragraphs with *both pro* and *con* annotations may have helped students work through conflicting perspectives on the materials and lead them to articulate a more nuanced and original opinion. After reading the two annotations (the first, *pro* annotation stated, "*I agree we need to stick to basic skills"* while the second *con* annotation stated, "*No! A liberal arts curricula could easily be integrated with computers*."), Tabetha comments " Umm, I think I agree with the second sidenote." However, rather than simply moving on after stating her opinion as she did many other times in the protocol, Tabetha continues her thought process:

> I think a *liberal arts curriculum* is important for what the article said, for developing students' values and intellect and for helping them think critically and being able to analyze information, but, umm, there is a way to *integrate computers* with that. I don't think it should be, umm, extensively computer-oriented, but I think that people do need a *basic*, umm, understanding of computers. *(bold indicates wording that repeats or references the primary text and underlining indicates wording that repeats or references the annotations.)*

The annotations thus seemed to prompt Tabetha to reflect both on the article's claims and her position on the issue. Such reflective moves in Tabetha's protocol—like those of most of the novice readers—were generally in response to annotations. Rarely did the primary text provoke such extended responses. Moreover, the underlining and bold in the quote above shows how Tabetha moved between primary text and annotation, reflecting on both as she moved towards articulating her own stance on the issue.

Other readers similarly transitioned between annotations and primary text in their responses.

Toby (novice reader):

> It says [rereads] **"This is one reason that school traditionalists push for broad liberal-arts curricula, which they feel develop students' values and intellect, instead of focusing on today's ideas about what tomorrow's jobs will be."** I think that [long pause] On that—on that, I go with—with D.H. It says: [rereads] *We need to stick to basic skills.* I think that *it could be easily integrated with computers*, but sh—the computers should just be a side note and shouldn't be, umm, *integrated the way that I think he's wanting* that, so, umm, I would say that the—the important thing in—in students—the important thing for **students to learn is to develop the values and intellect like they're talking about**. And definitely instead of—**instead of focusing on what tomorrow's jobs will be**.

Andrea (advanced reader):

> I agree with *the second* [annotation]. [writes and speaks]: **"Values and intellect"** could *easily be integrated with computers* [circles annotation]
>
> (*bold indicates wording that repeats or references the primary text and underlining indicates wording that repeats or references the annotations.*)

In this response, Toby moves back and forth between reading and reiterating the primary text and responding to the two annotations. Andrea similarly comes to a stance on this passage (which she writes down to return to when she drafts her essay) that integrates both primary text and annotation content. The ease with which these two readers transition between primary text and the two annotations seems greatly facilitated by their proximity to one another on the page. This proximity appears to help these readers reflect upon and integrate the conflicting viewpoints.

*Interviews*

In their final interviews, most readers commented that they found the annotations helpful and offered that the annotations influenced their reading. In particular, the participants commented that annotations that disagreed with the primary text were the most helpful. For instance, Eli (novice reader) commented:

> The ones that really helped me were when they'd disagreed with what was in the paragraph when they're like "Hey, this stuff's not necessarily true, how do you know it's true?" And the times that they did agree, then it didn't really help me that much cause I just read it and I'm like "Oh, that makes sense," and they're just basically telling me "Yeah"…. I like things that contrast. I think I can see them better when they contrast.

Other readers, however, clarified that the annotations that were most helpful were the ones *they personally* disagreed with. Toby (novice reader) responded:

> Only a couple of [annotations] I really disagreed with. But they're also helpful because they help you see another point of view that you wouldn't have thought of. I never would've considered that because I don't think it's an issue, but [those disagreements] led me to show another point by saying that that's not the case.

Jennifer (advanced reader) similarly commented, "If they agreed with me it wasn't validating or anything. I think it was more interesting if they disagreed because I could be like 'that's absurd' and it was more fuel for my argument." Tabetha (novice reader) also noted that the annotations she disagreed with could be helpful for understanding when she needed to concede to "the opposing side." Thus, for these readers, annotations that disagreed with their personal take on the issue seemed to make them realize that there was more to say and that additional concessions or defenses of their position were needed.

Perhaps because readers liked the contrasts, several noted that the paragraphs that contained both pro and con annotations were the most helpful of all. Andrea (advanced reader) clarifies:

> I liked when there were two annotations on the same things kind of with an opposing point because it made me think, well, obviously I can't just agree with it. I have to think would I side with this person or that person or which way do I feel about it.

Several readers also commented that the presence of both pro and con annotations made them undertake more effort to understand and address both sides of the issue. Eli (novice reader) states that it was "probably a lot because of these annotations…that I took both sides [in the response essay]." Wes (novice reader) similarly noted that "when the comments off to the side were different, it was really easy to look at both sides."

Not all of the interview comments on the annotations were positive. Jennifer (advanced reader) stated that the annotations "seemed really elementary actually and kind of silly. They just didn't seem very helpful," although when it was pointed out that she had circled and interacted with the annotations, she stated that she used them to continue her "dialogue" with the text: "to me like it was really just an extension of the essay where I could consider, I could just consider my argument even more." Kathy (novice reader) claimed that she would not have normally read the annotations but that she found while some were "pointless," others were "interesting" because they helped her see "a different point of view." Two other novice readers suggested that the think-aloud protocol study made them pay more attention to the annotations than they would have normally—although both claimed that they would have at least skimmed them under normal conditions. Only one participant (Toby, novice reader) suggested that the annotations could have interfered with his reading because "physically…I felt like I was jumping over there and I wasn't getting through the whole thought, so I ended up having to go back and reread it." However, it is quite possible that this interruption and consequent rereading of the primary text was actually a benefit to novice readers, causing them to pay more careful attention to the argumentation in the article than they would have otherwise. Overall, Toby claimed in his interview that the annotations were much more helpful than distracting.

## Study 2: Effects of annotated readings on student writing

A second, between-subjects study was conducted to see how annotations might affect students working in a more naturalistic setting than a think-aloud protocol. By employing a less labor-intensive methodology, this second study allows us to involve a greater number of participants working on a typical school task—reading assigned materials and writing a short essay in response. Moreover, the between-subjects approach, which involved giving two groups of students identical base content, helps ensure that differences in students'

responses are attributable to the annotation format and not to differences in the content of the primary text.

## Method

### *Participants*

Eighty-two college students enrolled in six sections of introductory college writing classes at an urban research university in the mid-western USA were randomly assigned to one of the two conditions. All but three students were freshmen. College freshman are used as study participants because research suggests that freshman writers lack effective strategies for making arguments from sources (Kaufer and Geisler 1989). All students were native speakers of English.

### *Study materials*

Two sets of materials were prepared. These two sets of materials were identical in content but differed in format. One group (the aligned annotation group) saw responses to the primary text in the form of annotations while the second group (the appended response group) saw the same response content in the form of informal essays appended to the end of the materials.

Both groups of students received the full excerpt of *The Computer Delusion* plus the letters to the editor described in the pre-study above. Students in the annotation condition saw a total of 30 annotations (half pro and half con, with five paragraphs containing sets of conflicting pro/con annotations together) on both the primary essay of *The Computer Delusion* and the accompanying letters to the editor. Students in the response condition did not see any annotations but instead received two short essays that strung together the annotation content into fictional response essays from two individuals. Thus, the only difference between the two sets of materials was the format in which the information was presented: annotations vs. separated responses.

### *Task prompt and study procedures*

Students were randomly given either the aligned annotation or appended response version of the materials. Students were asked to take the materials home, read them, and write a response essay that would be graded as a homework assignment. The essay prompt was purposefully left vague so as to elicit a range of possible responses from students and called for students to write "a response essay that takes a position on these materials." Students were additionally told to assume their readers were familiar with Oppenheimer's article and that a guest lecturer would appear the following class period at which time they would be invited to participate in a study.

Students showed up with their completed assignments at the next class meeting. A researcher then walked students through a questionnaire in class that asked students about the reading and writing strategies they had used in preparing their homework essays. Students were then told about the purpose of the study. Those interested in participating then completed demographic questionnaires and signed consent forms allowing their materials to be included in the study. The majority of students completing the homework assignment agreed to participate.

*Essay analysis*

Each essay was rated by three composition instructors on the following criteria: overall quality of argument, use of primary sources, use of compromise, and use of summary. A four-point Likert scale was used for each of these criteria and the scores of the three instructors were then averaged together. Interrater reliability between the three composition instructors ranged from low (average $r = 0.36$ for use of primary sources) to fair (average $r = 0.49$ for use of summary). Such low reliability is common when rating written products (Elbow 1993)—particularly for such an open-ended assignment—and is the reason that three raters were used for the scoring.

Table 3 describes the conditions for receiving a high score of four on each criterion. These criteria were chosen to target aspects of students' writing and argumentation that would most likely be influenced by annotations. In particular, "use of compromise" was chosen as a criteria because many of the participants in the think-aloud study noted that the annotations helped them understand both sides of the issue and made them make more of an effort to bridge these two sides. Similarly, "use of summary" was chosen because the annotations appeared to prompt the novice readers to move from purely comprehension strategies to more evaluative, reflective reading strategies. Such shifts in reading strategies may potentially be accompanied by writing strategies that focus less on summary (simple recounting of what was read) and more on persuasion (arguing for or against a particular position).

*Questionnaires*

After completing their essays as homework, students completed an in-class questionnaire asking them to identify the reading and writing strategies that they used in preparing their essays. Students received a series of descriptions of strategies and rated these strategies on a five-point Likert scale where 1 indicated "did not use at all" and 5 indicated "used this strategy heavily." This questionnaire was based upon previous research suggesting that novice college writers given reading-to-write assignments tend to rely on either summary strategies that reproduce the main arguments in the readings with little new information added or "springboard" strategies that focus on the students' own opinions on the topic with little reference to the readings. By contrast, more experienced writers tend to use strategies that will help them integrate, or bridge, their own opinions with those in the source readings (Kaufer and Geisler 1989). While self-reported data about writing strategies is suspect since people are often unable to accurately describe their thinking processes, it does offer some provisional clues as to what participants may have intended as they prepared to write.

**Table 3** Criteria instructors used to evaluate essays

| Criteria | Description of high score |
|---|---|
| Overall quality of argument | Makes a clear, original argument effectively supported by evidence |
| Use of primary sources | Analyses, evaluates or synthesizes primary sources to make an original observation about the materials |
| Use of compromise | Builds a bridge between competing positions, articulating a viable compromise that could appeal to both parties |
| Appropriate use of summary | Summarizes only to establish context; the amount of summary is equivalent to what might appear in a published response |

In addition, after turning in their essays and strategy questionnaires, students completed an additional demographic questionnaire. Students were also asked questions about the usefulness of the annotations or the appended responses, depending on which condition they were in.

## Results of between-subjects study

### *Self-reported reading-to-write strategies*

Table 4 shows that students receiving annotations described themselves as less reliant on summary strategies characteristic of novice writing. Students receiving annotations were less likely than those receiving appended responses to describe themselves both as using pure summary, $F(1,81) = 9.79$, $p < 0.01$ and simple summary plus opinion strategies, $F(1,81) = 7.64$, $p < 0.01$. These findings reinforce the think-aloud study which found that annotations encouraged students to move away from comprehension strategies towards more critical argumentative reading strategies. Thus, students receiving annotated materials were less likely than those in the other group to describe themselves as simply reproducing what they had learned.

However, Table 4 suggests that the annotations did not necessarily motivate students to use more advanced writing strategies in their essays. In fact, they were marginally less likely than students in the appended response condition to use one of the advanced strategies: evaluate the texts. This finding is challenging to interpret, but could be attributed to the fact that the appended responses were in essence short essays that evaluated the text. Thus, students in the appended response condition may have perceived these appended responses as model essays calling for textual evaluation.

### *Essays*

Not only did students in the annotation condition perceive themselves as less reliant on summary, but the instructors evaluating their essays similarly perceived them as less reliant on this novice strategy. The fact that such a small intervention influenced students' written products is somewhat remarkable when we consider that composition instructors often spend entire semesters attempting to get students to move beyond "memory dump" (Flower



**Table 4** Students' average self-reported use of reading-to-write strategies (and SD) on a five-point scale (1="not at all"; 5="strongly")

| Strategy | Aligned annotation | Appended response |
|---|---|---|
| Novice strategies | | |
| Summary only* | 2.03 (1.10) | 2.80 (1.14) |
| Summary plus opinion* | 2.95 (1.24) | 3.73 (1.30) |
| Springboard | 3.65 (1.31) | 3.26 (1.45) |
| Advanced strategies | | |
| Dig out an organizing idea | 2.20 (1.09) | 2.37 (1.24) |
| Interpret the text for own purpose | 2.70 (1.22) | 2.38 (1.39) |
| Evaluate the texts** | 2.23 (1.21) | 2.77 (1.28) |
| Mine the text for evidence | 2.75 (1.24) | 3.06 (1.36) |
| Do something for the reader | 2.10 (1.10) | 1.88 (0.95) |

*$p<0.01$, **$p<0.06$

JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008


1979) essays that simply reproduce what they have read. The bottom row of Table 5 indicates that students who received annotated materials were less likely to include inappropriate amounts of summary than students in the appended response condition, $F(1,81)=4.1$, $p<0.05$. This finding supports the results of the think-aloud study which similarly found that annotations made students less dependent on comprehension (i.e., summarizing and paraphrasing) reading strategies.

Moreover, Table 5 suggests that students who received annotated materials were judged by the three instructors as making marginally better use of compromise by trying to bridge competing positions, $F(1,81)=3.5$, $p<0.07$. While this result is not significant, it is suggestive given that one of the main benefits of annotations found in the think-aloud study was their ability to provoke students to consider the issue from a different point of view. Further research is needed to test if annotations that present conflicting viewpoints can stimulate readers to think through competing positions.

*Responses to the annotations*

Students' responses to the annotations on the final questionnaire support the comments made by participants in the think-aloud study. Students primarily commented that the annotations were helpful in seeing both sides of the issue and that the multiple perspectives encouraged them to reflect more thoroughly on the materials. One student, for instance, commented "I believe they prompted me to think more about the text, and to develop an opinion." Another student wrote, "I disagreed and agreed with the annotations which also helped me to understand the article." Yet a third student commented, "I wanted to argue with the person who's [sic] view differed from mine."

One interesting finding, however, is that approximately 25% of the students who received the annotations declined to read them. These students stated that they were in too much of a hurry to read the supplementary materials or else commented that they feared that the annotations would affect their viewpoint too much. This finding is discussed more in the implications.

## Discussion

The two studies here present complementary evidence suggesting that aligned annotations have considerable potential to shape students' reading practices. Both studies found that annotations encouraged students to move away from simple comprehension/summary response strategies and engage in more critical responses that reflect active, independent thinking. In this, they are joined by a third study (Wolfe and Neuwirth 2001), which used a

**Table 5** Average ratings the three instructors assigned to essays (4=high; 1=low)

| Criteria | Aligned annotation ($n$=40) | Appended response ($n$=42) |
|---|---|---|
| Quality of argument | 2.39 | 2.19 |
| Use of sources | 2.24 | 2.01 |
| Use of compromise** | 2.32 | 2.05 |
| Use of summary* | 3.13 | 2.75 |

*$p<0.05$, **$p<0.07$

different set of materials and similarly concluded that annotations move students away from summarizing towards more complex forms of engagement with the primary text.

Thus, it may be inferred from these studies that much of the learning value of an annotation lies in its ability to provoke students to take a stance on the primary text. Accordingly, we see from the think-aloud study that students were significantly more likely to engage in critical, independent thinking when reading annotated paragraphs, whereas their cognitive efforts were largely limited to comprehension (summarizing and paraphrasing) when reading non-annotated paragraphs. Complementing this result, the between-subjects essay study found that students receiving annotated materials both perceived themselves and were perceived by their instructors as less reliant on summary strategies than students who did not receive the annotations. These results are supplemented by an earlier study using different materials which also found that students working from annotated texts were less likely to produce "memory dump" essays of uncritical summary (Wolfe and Neuwirth 2001). While students in the annotation group did not produce significantly higher quality arguments overall, this may be because argument quality encompasses factors such as clarity of phrasing and organization that are unrelated to the type of critical thinking and perspective-taking annotations appeared to provoke.

In response to the question of what types of annotation have the most learning value for students, the think-aloud study indicates that annotations had the largest effect on students' critical thinking when readers encountered both positive and negative annotations on the same segment of primary text. The presence of two conflicting viewpoints in the margins appeared to stimulate readers to think through the issue, evaluate the claims more closely, and reflect more thoroughly on their own positions. In individual interviews, students particularly singled out as beneficial those places where conflicting annotations challenged them to consider and weigh "both sides" of the issue. As one student said, "I like things that contrast. I think I can see them better when they contrast" while another student noted that the conflicted perspectives required her to take a stance because "obviously I can't just agree with it." In light of these student comments, it may be worth noting that students receiving annotated materials were perceived as marginally more successful in building compromise positions that could appeal to individuals on both sides of the debate. Future research may want to examine the hypothesis that annotations reflecting competing positions could encourage readers to build compromises.

Critical reading similarly seemed to increase in paragraphs that contained a single *con* annotation disagreeing with the primary text, although the effect was not as strong. Participant interviews and prior research suggest that the ability of single con annotations to stimulate critical thought may depend on whether or not the annotation corroborated or conflicted with the reader's own pre-existing viewpoint on the materials. In general, students appreciated annotations that required them to consider "a different point of view." Prior research has found that the quality of students' essays seems to correspond with whether or not they agree with the annotations: when annotations simply support students' own perspectives they wrote lower quality essays, presumably because they saw no need to persuade their readers (Wolfe and Neuwirth 2001). Paragraphs with both pro and con annotations avoid this potential "preaching to the choir" drawback by ensuring that at least one of the annotations will contain content that the reader disagrees with.

While it is possible that the results in the think-aloud study could be attributed to the contents of the paragraphs themselves, this seems unlikely. First, care was taken to make the annotated and non-annotated materials equivalent. Second, if the differences in student strategies were attributable only to differences in the primary text, we would not see an effect in the between-subjects study which presented identical materials to students in

JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008


annotated and non-annotated format, thus strongly suggesting that the annotations—and not some other factor—were responsible for the shifts in reading strategies. Finally, the students themselves in both studies reported that they perceived the annotations as stimulating their critical thinking.

The paper-based materials used in this study were both an advantage and a drawback. The paper was an advantage because this study allowed testing of what may be the ideal conditions for displaying aligned annotations: annotations were relatively short (two sentences or less), precisely positioned to line up with the primary text, and anchored to precise segments of the primary text with underlining. The readers in this study seemed to take advantage of the precise positioning offered by the paper-based materials in order to jump back and forth between primary text and annotation. None of the participants in the think-aloud study evidenced any trouble in reconciling or matching up the primary text and annotation. Future research needs to assess whether aligned annotation formats that deviate from this ideal would produce similar results in students' critical reading activities.

## Pedagogical and design implications

1. Viewpoint of annotation is more important than quantity

As described in the introduction, one of the biggest problems with annotation systems is deciding how to manage screen real-estate when large numbers of annotations are present. One of the biggest contributions of this research study is implication that the viewpoint presented in the annotation had a major impact on reading practices. The presence of "pro" annotations that simply reaffirmed the primary text had little impact on reading practices while other types of annotations have significant effects on the quantity and quality of reflection students engaged in. Thus, this research shows that some annotations have more learning value than others. Wolfe and Neuwirth (2001) similarly suggests that some annotations may even have negative effects on students' argumentation skills. These findings indicate that CSCL developers should focus on designing annotation systems that present the best, most productive annotations rather than attempt the Sisyphean task of designing interfaces that can display all annotations made.

2. Display pairs of annotations that present conflicting perspectives on the primary text

This research found that the "best" annotations were those that provoked students to consider and weigh competing viewpoints on the primary text. This study indicates that the best way to inspire this type of perspective-taking is to provide students with pairs of annotations that contradict one another in the stance they take on the primary text. As the students themselves claimed, when the annotations presented conflicting perspectives, they were motivated to work harder to understand the text and reason through their own stance on the material. By contrast, when the annotations simply confirmed students' viewpoints, there was little effect on reading practices.

This finding raises the question of how to programmatically identify the annotation stance so that systems can filter for and select competing perspectives. One way to identify annotation stance is simply to have readers identify the stance of their annotation at the time the create it. Since only a small proportion of private annotations become public (Marshall and Brush 2004) it is possible that readers motivated to share an annotation would also be willing to identify its stance. A second possibility is to ask subsequent readers of the annotation to identify its stance. Instructors who are planning to assign a certain set of

materials to their students, for instance, might be motivated to contribute such meta-data in order to improve how the annotations are presented to students. A third method is to develop text-mining tools that can identify the stance of the annotation. For instance, words such as "however," "but," "I disagree" and "no!" are usually clear indicators of disagreement while phrases such as "Good point," "Yes," "agreed," and "exactly" indicate agreement. Many of the annotations produced by instructors in the pre-study began with such position words; therefore, it may be possible to develop a set of common words and their variations that can be used to identify an annotation's stance (see also Cabanac et al 2007 for related work).

One question that future research should consider is the extent to which the quality of expression or the validity of the viewpoint in the annotation matters. In other words, can even poorly phrased or reasoned annotations prompt student critical thinking as long as they present a new perspective for students to consider? The comments made by students in the interviews seems to indicate that the answer to this question would be "yes" since students remarked that even "silly" or "pointless" annotations encouraged them to expand their arguments and engage further with the text.

3. Students may benefit from instruction on how to use annotations

Approximately 25% of the students in the essay production study indicated that they had not read the annotations. Similarly, one of the five novice students in the think-aloud study stated she would not have read that annotations under normal circumstances. This suggests that students may need some coaching on how to use the annotations effectively.

Kathy, the student in the think-aloud study who said she would not normally have read the annotations, stated that she was used to seeing annotations in her high school textbooks. These textbook annotations were used to summarize key points that the authors did not want students to miss, and Kathy indicated that she typically ignored such authoritative annotations as redundant. Thus, her inclination to skip over the annotations in this study stemmed from her association with a different type of annotation in a different context. Kathy also indicated that she was unfamiliar with argumentative writing and had never before received an assignment that asked her to take a stance on a controversial reading. This lack of familiarity with argumentative writing (a common scenario for students just entering college) may have contributed to her inclination to ignore the annotations. Extrapolating from Kathy's experience, students in the between-subjects study may have skipped the annotations because they were unfamiliar with the perspective-taking required in argumentative writing and/or because they transferred their associations of annotations in a different context (textbooks) to this one (discussion).

Students, therefore, may need instruction in the benefits of examining issues from multiple perspectives and reassurance that rather than providing the "correct" answer (as textbook annotations often do), these argumentative annotations provide new material and points for reflection. This study indicates that for novice readers at least, annotations provide a scaffolding that prompts them to take a stance on the material rather than simply parroting back what has been said. However, students may need encouragement to achieve these benefits.

4. Annotation systems should be designed with reading practices in mind

Research on annotation in the digital library often focuses on supporting information retrieval, including the use of annotation to help readers locate relevant or high-quality texts (Agosti et al. 2006; Cabanac et al. 2007; Golovchinsky et al. 1999) or to develop effective summaries of texts (Marshall 1998). Other research focuses on how annotation systems can

JrnlID 11412_ArtID 9040_Proof# 1 - 26/02/2008



influence the types of comments readers produce, studying how the annotation format influences the attention readers pay to particular details of a text when responding to it (Brush et al. 2002) or how annotation systems can foster reading communities and social interactions (Kaplan and Chisk 2005). Relatively few studies, by contrast, have looked at how the presence of annotations left on a document by previous readers can influence an individual's perception of that document and their experience of its claims.

This study indicates that readers have a different relationship with annotations than they do with the primary text. Where novice readers adopted a generally passive reading style when reading the primary text, they were more proactive and questioning when reading the annotations, which seemed to provoke them to take a position. This may be because the readers were mimicking the reading practices modeled in the annotation—or it could be that the proximity between the annotation and the main claim in the primary text helped reduce the cognitive effort required to balance and weigh competing perspectives.

In either case, this shift in reading strategies can be useful for scaffolding critical thinking practices in educational settings where instructors often want students to wrestle with complicated and controversial issues. As educators design e-learning sites for collaboration and discussion, they would do well to keep in mind how even simple changes in technologies for reading and writing have the potential to shape the activities and learning that go on in these places.

Designers interested in creating sites for civic participation and debate should also look carefully at how the display and organization of responses could affect the quality and tenor of discussion. Newspaper websites, for instance, often have discussion forums and published "letters" pages where readers can discuss articles or editorials. Designers of these sites might improve the quality and potential usefulness of these discussions by switching to annotated displays that could facilitate subsequent readers in assimilating the competing perspectives presented. These annotated displays might thus encourage readers of these opinion pages to move away from simple "pro" or "con" reactions to debates and toward a more integrated reconciliation of competing viewpoints.

As the Digital libraries, persistent conversation, CSCL, and other research communities move towards increasingly advanced systems for capturing, storing and displaying annotations, one measure of the effectiveness of these systems should be how they can improve upon readers' abilities to consider, reflect on and weigh competing viewpoints. We need therefore to keep in mind not just the needs of speakers—those who make comments—but also the needs of listeners, those who read what others have contributed. Ultimately, it may be the listener/readers who have the most to gain.



**Acknowledgement** This research was supported by a University of Louisville Intramural Research Incentive Grant. The author would like to thank Andrea Ascuena for her help in coding the think-aloud protocols and the many students who participated in the study.

## References

Agosti, M., Ferro, N., Panizzi, E., & Trinchese, R. (2006). Annotation as a support to user interaction for content enhancement in digital libraries.*Proceedings of the working conference on Advanced visual interfaces*. Venezia, Italy: ACM.

Brush, A., Bageron, D., Grudin, J., Borning, A., & Gupta, A. (2002). *Supporting interaction outside of class: Anchored disccusionss vs. Discussion boards.* Paper presented at the Proceedings of CSCL 2002.

Cabanac, G., Chevalier, M., Chrisment, C., & Julien, C. (2007). Collective annotation: Perspectives for information retrieval improvement, *RIAO 2007*. Pittsburgh, PA.

Charney, D. (1993). A study in rhetorical reading: How evolutionists read "the spandrels of san marco". In J. Selzer (Ed.) *Understanding scientific prose* (pp. 203–230). Madison: University of Wisconsin Press.

Elbow, P. (1993). Ranking, evaluating, and liking: Sorting out three forms of judgment. *College English*, *55* (2), 187–206.

Flower, L. (1979). Writer-based prose: A cognitive basis for problems in writing. *College English*, *41*(1), 19–37.

Glass, G. (2005). Marginalia. Retrieved June 20, 2007, from http://www.geof.net/code/annotation/demo/.

Golovchinsky, G., Price, M. N., & Schilit, B. N. (1999). From reading to retrieval: Freeform ink annotations as queries. In *Proceedings of sigir '99* (pp. 19–25). New York: ACM.

Guzdial, M., & Turns, J. (2000). Effective discussion through a computer-mediated anchored forum. *Journal of the Learning Sciences*, *9*(4), 437–469.

Honneycutt, L. (2001). Comparing email and synchronous conferencing in online peer response. *Written Communication*, *18*(1), 26–60.

Jackson, H. J. (2001). *Marginalia: Readers writing in books*. New Haven: Yale University Press.

Kaplan, N., & Chisk, Y. (2005). In the company of readers: The digital library book as "practiced place". In *Proceedings of the joint conference of digital libraries '05* (pp. 235–244). New York: ACM.

Kaufer, D. S., & Geisler, C. (1989). Novelty in academic writing. *Written Communication*, *6*(3), 286–311.

Kienle, A. (2006). Integration of knowledge management and collaborative learning by technical supported communication processes. *Education and Information Technologies*, *11*(2), 161–185.

Lebow, D., & Lick, D. (2002). *Hylighting: A new tool for distance and distributed teaching and learning, The Eighth Sloan-C International Conference on Online Learning*. Orlando, FL.

Lunsford, A. A., & Ruszkiewicz, J. J. (1999). *The presence of others* (3rd ed.). New York: St. Martin's Press.

Marshall, C. C. (1997). Annotation: From paper books to the digital library. In *Diginal libraries 1997* (pp. 131–140). New York: ACM.

Marshall, C. C. (1998). Toward an ecology of hypertext annotation. In *Hypertext98* (pp. 40–49). Pittsburgh, PA: ACM.

Marshall, C. C., & Brush, A. J. B. (2004). Supporting personalization: Exploring the relationship between personal and public annotations. In *Proceedings of the 2004 joint ACM/ieee conference on digital libraries* (pp. 349–357). Pittsburgh, PA: ACM.

Oppenheimer, T. (1997/1999). The computer delusion. In A. A. Lunsford, & J. J. Ruszkiewicz (Eds.) *The presence of others* (pp. 255–285). Boston: St. Martin's Press.

Ovsiannikov, I. A., Arbib, M. A., & McNeil, T. H. (1999). Annotation technology. *International Journal of Human-Computer Studies*, *50*, 329–362.

van der Pol, J., Admiraal, W., & Simons, P. R. J. (2006). The affordance of anchored discussion for the collaborative processing of academic texts. *International Journal of Computer Supported Collaborative Learning*, *1*(3), 339–357.

van Oostendorp, H. (1996). Studying and annotating electronic text. In J.-F. Rouet, J. J. Levonen, A. Dillon, & R. J. Spiro (Eds.) *Hypertext and cognition* (pp. 137–148). Mahwah, NJ: Lawrence Erlbaum Associates.

Wojahn, P. G., Neuwirth, C. M., & Bullock, B. (1998). Effects of interfaces for annotation on communication in a collaborative task. In *Chi '98, conference proceedings on human factors in computing systems* (pp. 456–463). New York, NY: ACM.

Wolfe, J. L., & Neuwirth, C. M. (2001). From the margins to the center: The future of annotation. *Journal of Business and Technical Communication*, *15*(3), 333–371.

Zellweger, P. T., Regli, S. H., Mackinlay, J. D., & Chang, B.-W. (2000). The impact of fluid documents on reading and browsing: An observational study. In *Proceedings of the chi 2000 conference on computer-human interaction* (pp. 249–256). New York, NY: ACM.