# From Personal to Shared Annotations

**Catherine C. Marshall**
Microsoft Corporation
cathymar@microsoft.com

**A.J. Bernheim Brush**
University of Washington
ajb@cs.washington.edu

## ABSTRACT
Preliminary results obtained by comparing personal annotations on paper with shared annotations made on-line show that only a small fraction of personal annotations are used in initiating and responding to related on-line discussions. The personal annotations that are shared tended to correspond to explicit marginalia; much effort is still put into rendering both the content and anchors of these annotations intelligible to others.

## Keywords
Annotation, collaboration, reading, on-line discussion

## INTRODUCTION
People often annotate paper documents as they read them, especially if they are responsible for assimilating the content [3]. They underline text, write notes in the margins, place asterisks by content they want to find again, and otherwise create a personal geography of the reading materials. Currently, if these annotations are shared, the sharing is serendipitous, for example through the circulation of photocopied paper documents at work or through the sale of used textbooks in education [2].

With the advent of better computer support for reading and annotating documents, we speculate that more personal annotation will be done on the computer. What will new technologies mean for the sharing of personal annotations? Just as it will be easier to pass along personal annotations, it will also be easier to remove them. Which annotations will people want to intentionally share? Will other readers be able to make sense of these annotations?

Thus we are interested in how personal annotations are used in collaborative situations and the transitions the annotations undergo when they are shared with others. Are there patterns that will allow us to predict which of the annotations will be useful and how they will be shared?

Our hypotheses about the relationship between personal and shared annotations are: (1) that only a small fraction of the informal personal annotations are ever shared with others; (2) that some types are more likely to be shared than others; (3) that annotations made while reading undergo a profound transition when they are used as part of an on-line discussion. As part of this transition, anchors are made more coherent and may either grow or shrink to better define the topic under discussion. More importantly, much effort goes in to rendering the content of personal annotations intelligible to others.

## METHOD
To understand the relationship between personal annotations people make while they are reading and the annotations they contribute in a collaborative setting, we compared graduate students' personal and on-line collaborative annotations for a graduate-level computer science course. The assigned on-line discussions took place using WebAnn [1] a Web browser based tool that supports discussions anchored directly to the source material and displayed in context. The eleven students used WebAnn for 4 (non-consecutive) weeks to discuss twelve publications.

Figure 1 shows a student annotation on a printed article, and the discussion she initiated in WebAnn that corresponds to her personal annotation. In WebAnn anchors are indicated by boxed text in the source HTML documents; the remarks themselves are shown in a separate pane on the side of the browser. Users can navigate readily between the annotations and the source text they refer to; thus the annotations are comparable to a "highlight and margin note" style of commenting. Users can also initiate new discussions and reply to existing annotations.
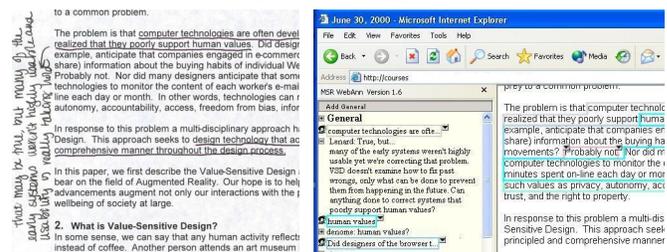


**Figure 1. A personal annotation on paper and its shared on-line analog in WebAnn**

To obtain the preliminary results that we present in this paper, we worked with a representative reading from each week the students used WebAnn and collected 6, 6, 5, and 3 papers annotated by students for the four weeks. (Out of the 11 students, 2 made no written comments and 3 were not able to give us their papers for analysis.) Note as the term went on more students used the option to skip some of the readings, so fewer papers were available. As part of our future work, we intend to perform a more detailed analysis on the remaining eight sets of readings.

We coded the student's annotations on paper and in on-line discussions to record the type and content of each annotation, its anchor type, any correspondences between

personal and collaborative annotations, and how the anchors and content had changed in the event of a correspondence. For annotation type, we captured fairly fine-grained distinctions (for example, between a highlight and an underline), and recorded whether the annotation was compound (an underline and margin note, for example). For anchor type, we coded the extent of the anchor (e.g. whether it spanned a word, a phrase, a sentence, multiple sentences, or an entire paragraph).

To ensure uniform application of the coding scheme to borderline cases, both authors coded the annotations independently and compared their results. We also compared our interpretations of whether the personal and collaborative annotations corresponded to one another.

In addition to holding on-line discussions in WebAnn, the students were required to write a separate summary in the system. We recorded which personal annotations contributed to this summary. This way, as part of our investigation, we can explore whether personal annotations have a greater tendency to contribute to the summary the student wrote to describe the paper, or the discussion itself.

## RESULTS
As part of our initial analysis, we characterized the kinds of personal annotations the students made on paper, and how they compared to the collaborative annotations they made on-line. Out of a total of 602 annotations, 504 (84%) were on paper and 98 (16%) were on-line; of the 98 on-line annotations, 20 were the required summaries, and the other 78 were part of the on-line discussion.

Of the personal annotations made on paper, 414 (82.1%) were simply underlines, highlights, and circles (i.e. well-specified stand-alone anchors, with the majority, 302, being underlines), 38 (7.5%) were compound (an anchor combined with a note or symbol), and 23 (4.6%) were just notes (commentary without an anchor). The non-summary on-line annotations were all anchored notes (this is what the system supported); of these, 43 (43.9%) were notes on the readings and 36 (36.7%) were replies to others' notes.

Anchor types varied more on paper than they did on-line, although the two most common anchor types on paper, single sentences (152 anchors, 30.2% of the total) and phrases (124, 24.6%) were also the most common on-line, 42 (54%) and 13 (17%), respectively.

This preliminary data shows that the students made far more personal annotations than they did collaborative annotations. The personal annotations were more cryptic, mostly anchors without notes. The on-line annotations were also more likely to be anchored in complete sentences.

Given this broad-brush characterization, it is interesting to more closely inspect how the individual annotations on paper relate to those on-line. Overall, we found 137 correspondences between the annotations on paper and those on-line. Many of these correspondences, 96 (70%), were between personal annotations and the on-line summary. However, 35 of the matches demonstrated

relationships between personal annotations and online discussions. In 6 cases, the correspondence was between a personal annotation and both summary statements and commentary.

The small number of correspondences to online discussions shows that relatively few personal annotations contribute directly to the students' on-line commentary. Table 1 illustrates that although personal annotations with content (e.g. notes) occur infrequently on paper they are far more likely to form the basis of on-line commentary.

| Annotation Type | Frequency on paper | Number that correspond |
|---|---|---|
| Compound (anchor + note/symbol) | 38 (7.5%) | 15 (42.9%) |
| Note (unanchored) | 23 (4.6%) | 6 (17.1%) |
| Underline/highlight/circle | 414 (82.1%) | 14 (40.0%) |

**Table 1. Types of personal annotations likely to be shared**

Finally, we examine the transformations the personal annotations on paper undergo as they are shared on-line. The data show that 74.3% of the annotation anchors change; they are "rationalized" to more precisely delineate the annotation's scope. The data also show that the content of annotations change significantly. Only 14.3% are simply "cleaned up" versions of the notes made on paper. The remainder are substantially expanded and changed.

## DISCUSSION AND FUTURE WORK
Our preliminary results and observations lead us in the following directions: Will our results hold over the larger set of papers we have collected? Are there important differences in the annotation styles and behavior of particular students? What is the role of the students' annotations in the summaries they produce? What effect does an author's writing style have on a reader's annotation style? Finally, we plan to explore whether the annotations are a good predictor of which portions of the document would be discussed. Were points annotated by multiple readers more likely to generate discussions with responses?

The relationship between personal annotations and the document summary seems like a promising direction to explore based on our preliminary results (70% of annotations that were used on-line corresponded to summaries); personal annotations may be used to summarize a document more often than they are used as a basis for discussion.

## REFERENCES
1. Brush, A., Bargeron, D., Grudin, J., Borning, A. and Gupta, A. Supporting Interaction Outside of Class, in *Proc. CSCL '02* (Boulder CO, January 2002), 425-434.

2. Marshall, C.C. Annotation: from Paper Books to the Digital Library, *Proc. of DL'97*, ACM Press, 131-140.

3. O'Hara, K., Smith, F., Newman, W., & Sellen, A. Student Readers' Use of Library Documents, in *Proc. of CHI '98*, ACM Press, 233 – 240.