

On-Line Reviewing with Change Representation Tools

Hee-Cheol (Ezra) Kim

Department of Numerical Analysis and Computer Science
Royal Institute of Technology
S-100 44 Stockholm
Sweden

+46 8 790 62 81
hecki@nada.kth.se

ABSTRACT

Change representation concerns how to present changes of text in writing tools. The Reviewer-Initiative (RI) model of collaborative writing means that the reviewer revises on-line documents as suggestion of further revision, and the writer later either accepts or rejects each suggested change. This paper reports a laboratory study that investigates how people use tools for change representation under the RI setting. Reviewers made most revisions at a low level. Paper documents were not used so often. They felt that the commenting function should be used to complement the limitation of suggestions by change representation alone. Sense of hurting or being hurt, which can be caused by direct on-line revision, was little. Reviewers' difficulty of formulating words was not a great burden, and the process of accepting or rejecting each suggested change was not a hard decision-making to the writers. There was no significant difference between the RI on-line reviewing model and the traditional one with respect to writers' sense of help from reviewers, and reviewers' understanding of text.

Keywords

Change representation, collaborative writing, commenting function, on-line document, paper, reviewing

1. INTRODUCTION

During the last decade, computer-supported collaborative writing has gained growing interest from researchers and practitioners (Ede and Lundford 1990, Lay and Karis 1992, Rada 1996, Santos 1995, Sharples 1993). As a matter of fact, many professionals spending much time writing work collaboratively in various writing projects (Ede and Lundford 1990).

An important factor to get success in collaborative writing projects is the communication about changes of the document among the collaborators. Co-authors should understand what changes were made in the text in order to see the improvement or the difference of the current version from the previous text. The collaborators

constantly communicate to build the common ground on changes in the text and further planning of changes.

Therefore, one of interesting issues that we should consider is how the computer can help the collaborating writers understand and communicate about changes of text. Wide-spread writing tools such as MSWord and FrameMaker currently provide features to present changes of text, which we call *change representation* functions. Some users are already using these functions in certain co-authoring projects (Kim and Severinson Eklundh 1998). In our study, the notion of collaborative writing is not only confined to the cases where both a writer and a reviewer have co-authorship. Taking its notion more broadly in our work, we assume that the reviewer does not mean an author of the document, but

instead provides the writer with comments and possible revisions as a helper of the writer.

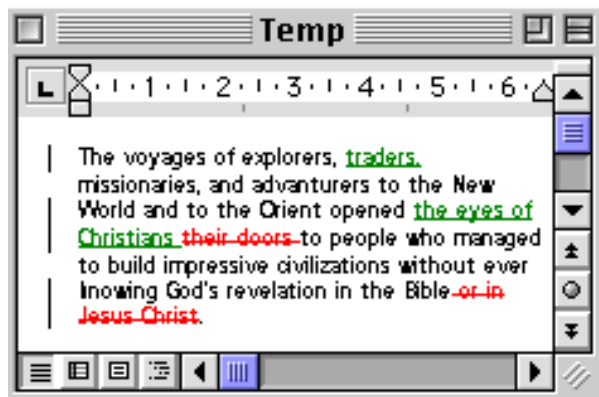


Figure 1. An example of the text with change representation in MSWord

In early 90s, some design issues in change representation were addressed by Neuwirth et. al. such as what changes to show on electronic documents and how to pinpoint changes (1992). Since then, however, change representation has not been high on the research agenda, while the tools supporting it have been rapidly developed. Against such a gap of awareness between the research community and the real world, we have previously conducted an experimental study to explore how people make use of change representation functions and draw implications of the design of the functions (Kim and Severinson Eklundh, 2000). In the study, different conditions were set: two ways of representing changes, indication and display, and two models of collaborating, the writer- and reviewer-initiative models.

1.1 Indication versus Display

Indication of changes means that a change representation tool visualizes only the locations where changes were made in the text, while display of changes refers to its support to show both locations and deleted/added texts. According to the results of the study (Kim and Severinson Eklundh, 2000), people prefer the indication mode for reading the text, and the display mode for understanding the rationale of changes. There was also a great difference of understanding changes between the two modes.

1.2 Writer-Initiative (WI) and Reviewer-Initiative (RI) Models

In the writer-initiative model, the reviewer suggests changes of text by making comments, without revising the on-line text directly. Then the writer revises the text based on the reviewer's comments. After the writer completes the revision process, the revised version of the text is passed to the reviewer. At this time, the reviewer can see the changes made to the text by the use of change representation tools.

In the reviewer-initiative model, on the other hand, the reviewer directly changes the on-line document rather than giving suggestions of changes by comments, using change representation tools. After receiving the file of the revised text by the reviewer, the writer either accepts or rejects each change. MS Word 97 or above for PCs has a special function called "Accept or Reject Changes" to allow such a process, for short the A/R function.

The previous study showed that change representation functions were more useful for the reviewer in the WI model than for the writer, while they were mutually beneficial in the RI model. The RI model especially led the collaborators to use the commenting function more.

1.3 Research questions

Reviewers are more accustomed to making comments on others' text, rather than changing their on-line documents directly (Kim and Severinson Eklundh 1998). However, considering a potential of the RI collaboration with change representation tools as an alternative to the current reviewing practice, this study was designed for further exploration of the RI model, with the purpose to examine several issues as follows:

- 1 What kinds of revisions are made by the reviewers?
- 2 Why does the reviewer use the commenting function and what kinds of comments are made?
- 3 To what extent do users make use of a paper copy of the document while reviewing?
- 4 To what extent are there emotional effects when the reviewer changes the on-line text directly?
- 5 How difficult is it for the reviewer to formulate exact words? How difficult is it for the writer to decide either to accept or reject each revision suggested?
- 6 Can the RI collaboration affect the reviewing time, the discussion at the meeting, reviewers' understanding of text, and writers' sense of help from reviewers?

2. DATA COLLECTION

We carried out a laboratory task where 13 pairs of first-year university students from our department were involved. Every freshman at the department has to pursue a mandatory course called "Communication" through which they might acquire basic knowledge and skills about how to write a scientific paper and how to communicate with people by writing. Most of them are somewhat inexperienced in writing a scientific paper.

At the final stage of the course, a laboratory assignment was designed such that each student wrote an essay about the topics that he or she selected, and about five pages long. Pairs of students, say 'S1' and 'S2', made up a group and they reviewed each other's draft. During the reviewing process, they worked by the RI collaboration model with respect to S1's document, while following the traditional way of reviewing S2's document. In case of the RI working, the reviewer was also allowed to use the commenting function at will. The traditional model refers to reviewing with pen and paper, that is, the reviewer writes comments on the paper copy of the document by pen.

There are two main sources of the data that we have obtained: (1) subjects' answers to the questionnaire distributed after the experimental task was completed, and (2) their documents such as the first draft, the revised version as a result of the reviewer's revision using the change representation function, and the document made after the writer accepted or rejected revisions suggested. The more detailed procedure of the lab is described below:

Step 1: Students S1 and S2 wrote their essay drafts, and handed in the file to their partner so that they could review each other's drafts. They used MSWord 97 or above on PC to perform the lab assignment.

Step 2 (RI model): S2 reviewed S1's document by the use of the change representation function (and the commenting function optionally) so that all the changes suggested might be recorded and shown on the screen. Five days were allowed to review the draft. After completing reviewing, S2 sent the file of the revised document to the original writer S1. Student S1 then accepted or rejected the changes proposed by using the A/R function. Two days were allowed for this.

Step 2* (Traditional model): S1 reviewed S2's document, commenting on the hardcopy with a pen. A week was given to do this.

Step 3: S1 and S2 had a face to face meeting to discuss each other's documents for changes of text and further planning. The meeting took about an hour.

Step 4: S1 and S2 filled out the questionnaire.

3. SELECTED RESULTS

3.1 Patterns of suggested revisions

The majority of suggested revisions from the reviewers were low level revisions (73.5%) and middle level revisions (25.5%). We only found two sentence level revisions (1.0%). This fact tells us that student reviewers are likely to suggest small detailed changes by the change representation function. They actually utilized the commenting function and the physical meetings to discuss higher level revisions.

How do writers reacted to suggested revisions? Writers seem to have little trouble to accept character level revisions made by the reviewers. In fact, they accepted all those revisions as shown in Table 1. Since character level revisions are mainly typing error and grammar error corrections, it was an easy decision for the writers to accept those revisions. However, when it comes to a word or higher level, the rejection rate grows (10.5% at one word revision, 21.6% at middle and high level revision). Two writers who encountered sentence level revisions reported their difficulty in deciding to accept or reject them. Another writer also mentioned that he several times hesitated between accepting and rejecting word level revisions.

	Low level revision			Middle level revision *	High level revision **	Total
	Punctuation	Character level revision	One word revision			
Suggested revision	21	63	57	49	2	192
Rejected revision	7	0	6	10	1	24

*Middle level revision: revisions of more than one word and less than a sentence. **High level revision means sentence level revisions.

Table 1. Patterns of suggested and rejected revisions

3.2 Usage of paper

Reviewing documents demands reading substantially. Reading on paper has numerous advantages. Reading from paper is faster than reading on-line (Gould et. al. 1987). Paper provides flexible annotation while reading, quick navigation from page to page, and better

understanding of text (O'Hara and Sellen 1997). It also offers readers a better sense of the overview of text (Severinson Eklundh 1992). As a matter of fact, paper is extensively used when people review documents

On the other hand, our electronic text environment is also transforming. We now have functions to present changes of text on screen and write comments on on-line documents. Such a new environment has other advantages which paper cannot provide such as visualization of changes, easy acceptance or rejection of changes, and storing comments in electronic documents. With the advance of technology, we might question what it will be to use pen and paper. Will current practices persist? Or will these change? In this study, focusing on the RI on-line reviewing, we would like to find an answer to the following question:

To what extent do the collaborators read paper documents in the RI setting?

With respect to this question, the reviewers were asked to answer when they read paper documents in terms of "before" and/or "while" they revised the text on-line by the change representation function. In the same way, the writers were also requested to answer this question in terms of "before" and/or "while" they accepted or rejected the suggested revisions by the A/R function. We hypothesized that most subjects would read paper documents as well as on-line documents, since reviewing with paper is quite common in co-authoring projects (Kim and Severinson Eklundh 1998). As we see it in Table 2, however, three reviewers used paper for reading, and one reviewer did it only before making changes. Paper was even less used to the writers, that is, one writer alone used it only while accepting or rejecting proposed revisions.

	Only before reviewing on-line	Only while reviewing on-line	Before and while reviewing on-line	No use of paper documents
Reviewers	1	0	3	9
Writers	0	1	0	12

Table 2. Number of use of paper documents

This result suggests that on-line reviewing in the RI setting can make people use paper documents less than other traditional ways of reviewing. Although we are cautious to generalize this result since (co-)writing is

very context-dependent, we believe that the use of paper can decrease in certain conditions like the RI on-line reviewing.

3.3 Usage of commenting function

The commenting function is indispensable for on-line reviewing. Though its choice was optional in the task, most reviewers in the RI setting (11 of 13) had used the commenting function to complement the use of change representation alone. In response to a question "Why do you think you have used the commenting function?", those 11 reviewers answered as follows:

Because it was better to write comments in some cases than to formulate exact words (11 reviewers)

Because I wanted to present both comments and the changes I suggested in an electronic form rather than to present comments with a pen on the paper copy of the document (6 reviewers)

Because I was curious about a new function, the commenting function (2 reviewers)

Reviewers mainly used the commenting function when there was a difficulty to formulate changes of text. Interestingly, some of them (6 of 11) also wanted to provide the writers with the comments as well as their suggestions of changes by one medium only.

Use of change representation alone lacks in conveying what reviewers would like to suggest exactly, and the commenting function is a good reinforcement for this in on-line reviewing. Therefore, it is interesting to investigate what kinds of comments reviewers made on their partners' documents. The 11 reviewers who used the commenting function made 64 comments on their partners' documents. During the analysis, the majority of those comments (46 of 64) have been categorized into the following two cases in terms of the difficulty of formulating exact words:

Category 1: no suggestion.

There were 21 cases (32.8%) that reviewers found problems in certain parts of the text, but could not suggest properly how to revise it. For example, "You would better change the word here", "knowledge about knowledge ... sounds a bit strange", "Unclear?", "Perhaps too much of your own point".

Category 2: slight suggestion

We define slight suggestion as the act of suggesting changes through the commenting function rather than the change representation tools. There seem to be

various reasons to use the commenting function to propose possible changes. Reviewers sometimes had weak confidence in revising the on-line document directly, saying e.g. "It can be changed to arbetare", "Perhaps neurala nätverk?" They also encountered repeated problems in the text, expressing e.g. "Too many '<>'s might be skipped". In some cases, reviewers could not suggest the detailed changes by the change representation function, saying e.g. "It would be good to write the report under one title, not giving different headings." We found 25 cases (39.1%) as slight suggestion.

Apart from these two main categories, 10 comments (15.6%) also reflected reviewers' poor understanding of some parts of the document (don't-understand): "It was hard to understand what you meant. Could be written in more detail", "Do you mean KB systems would be one of the systems used in real world?" We also found three comments (4.7%) saying the reasons why reviewers made revisions in the way where changes were presented on the on-line document (change explanation). The results are summarized in Figure 2.

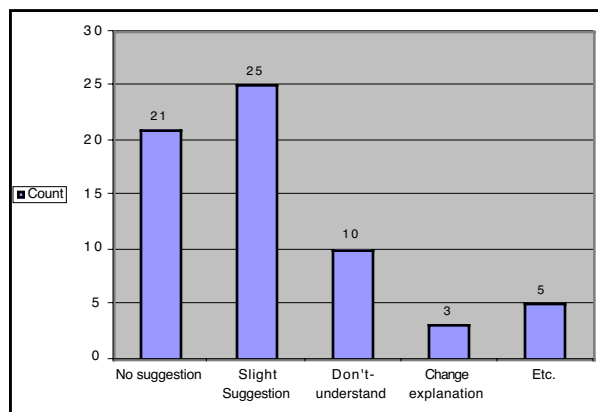


Figure 2. Patterns of comments

A fairly small number of comments for change explanation, however, was unexpected. Why did they make little use of the commenting function to account for their revisions? First, we believe that they did not need to do it, because most of the revisions that they made were obvious and low level revisions. Second, they perhaps want to explain revisions that they suggested, by using richer media like face-to-face which is more adequate for dealing with ambiguous and delicate matters. Third, they also might feel some cognitive overload to write comments for explaining the changes.

3.4 Affective aspects

In every kind of collaboration, human relationship between the collaborators is an important factor to bring success of their co-work. Emotion is one of crucial aspects to build their relationship. In writing, reviewers' suggestions that are difficult to accept from the writer's point of view can be a source of conflicts. Direct revision of the on-line document of the original writer's can be even more sensitive than provision of comments only, if it means the reviewer's ignorance of the writer's intentions and knowledge. It is thus interesting to see the extent of emotional effects in the RI model.

In response to the question "To what extent did the changes of your text made by reviewers hurt your feeling?", the writers were asked to rate 5 statements from 1 ("not at all") to 5 ("a lot"). 12 of all 13 respondents answered "Not at all", while only one writer gave 4 to the question. In the similar way, reviewers were asked to answer the question "When you change the text of the writer directly, how often did you think you might hurt the feeling of the writer?" The majority of them did not think of it seriously (8 for "never", 4 for "a few times"), whereas the response of one reviewer alone was "often". These results indicate that direct change did not affect the writers' emotions badly. When co-authors work together according to the RI model, we therefore think that feeling hurt concerns social aspects rather than the technology.

As patterns of revision were shown in the Table 1 before, reviewers tended to make revisions at a low level more than at the middle and the high levels. Since controversial parts of the document usually occur at the high level, reviewers may rely on the commenting function or physical meetings to discuss those parts so that they could avoid possible conflicts in advance caused by their revisions. In addition, mutual consensus between the writer and the reviewer on adopting the RI collaboration can also make the writer's feeling hurt about the revisions insignificant.

3.5 Difficulty of formulation and decision-making

The reviewers in the RI condition were asked to rate the degree of their difficulty of formulating exact words for suggesting changes of the writer's text, given 5 cases from 1 ("very easy") to 5 ("very difficult"). We anticipated that the reviewers would have a certain level of difficulty doing this, assuming that revision requires cognitive overload to some extent. However, 61.5% (8 of 13) out of the reviewers, which is a relatively high

percentage, thought it was (very) easy. We believe that this phenomenon occurs because reviewers are likely to change the parts on which they make revisions easily, only giving comments on difficult parts with the commenting function.

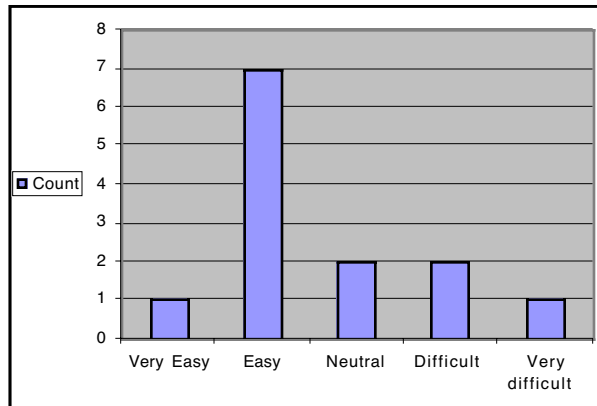


Figure 3. Perceived difficulty of formulating words

Writers in the RI condition were also asked to indicate the frequency of cases of their hesitation in terms of making a decision between accepting or rejecting each change made by the reviewer. The data shows that most writers (10 of 13) did not have much trouble to decide it (4 for "never" and 6 for "a few times").

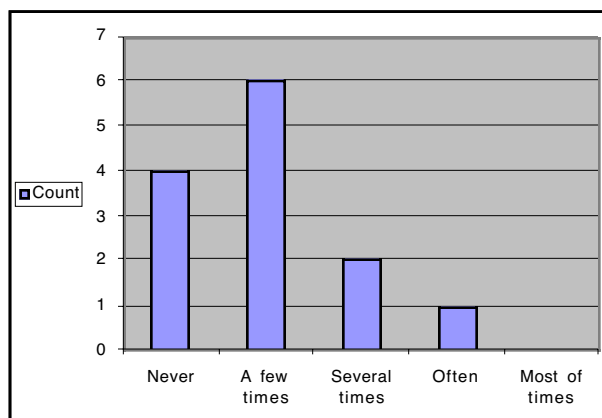


Figure 4. Perceived hesitation of either accepting or rejecting revisions

3.6 RI on-line reviewing versus traditional reviewing

Finally, we show the result of our investigation into if there are significant differences between the RI on-line reviewing and the traditional reviewing with respect to the following aspects.

3.6.1 Spending time

How long did reviewers spend in reviewing the document? Note that the result does not come from a certain accurate measure, but is dependent on subjects' own gauges, i.e. their answers appeared in the questionnaire. Writers produced 1,013 words on average per each document in the traditional model, and 864 words in the RI model. Reviewers spent 55 minutes on average in reviewing in the traditional setting, while it took 62 minutes in the RI setting. Therefore, the average time spent for reviewing per 1,000 words is 54.3 minutes in the traditional model, while it is 72.7 minutes in the RI model. After all, reviewers in the RI setting spent more time (33.9%) in reviewing than in the traditional setting.

3.6.2 Discussion at the meeting

Subjects were asked to answer how much they spent in discussing their documents at the meetings in terms of character or word, sentence, paragraph levels, and the whole document.

	Character or word level (%)	Sentence level (%)	Paragraph level (%)	Whole document (%)
Traditional model	8.3	16.7	22.2	52.8
RI model	9.0	18.0	22.5	50.5

Table 3. Percentage of the time spent at the meeting

The expectation that we had concerning the question was that more discussion at higher levels in the RI condition would be promoted. This was because we believed that groups in the RI model would focus more on higher level changes, since writers in the setting already accepted or rejected low level changes suggested by reviewers before the meeting. However, the data shows that there is no significant difference between the two conditions, as we see it in the Table 3 (9 groups

answered in the traditional setting, and 10 groups in the RI setting).

3.6.3 Reviewer's understanding of text

Can the RI model affect reviewers' understanding of the text? Reviewers in the RI condition spent more time in reviewing. It perhaps means that they thought about the text more than reviewers in the traditional way. If so, it is also possible that reviewers' understanding of text in the RI model is better than the traditional model. However, the result says that there is no great difference. In response to the question "To what extent do you think the reviewer understands your text?", writers' average rating was 2.6 in the traditional model, and 2.7 in the RI model on a 5-point scale (1 = "Not at all"; 5 = "A lot"). Also, reviewers own perception of understanding the partner's text was 2.5 and 2.8, respectively.

	Writer's perception	Reviewer's own perception
Traditional model	2.6	2.5
RI model	2.7	2.8

Table 4. Reviewer's understanding of text

3.6.4 Sense of improvement of text

How helpful are reviewers' suggestions in the two different conditions? We asked writers to reply to the question, "To what extent do you think the partner's suggestions helped you improve your document?", on a 5 point scale (1 = "Not at all"; 5 = "A lot"). The average values were 3.3 in the traditional model and 3.2 in the RI model, respectively. Therefore, we can see that there is no difference between the two conditions in assisting writers' text improvement.

3.6.5 Possibility of co-authoring without face-to-face communication

In response to the statement "Without face-to-face meetings, the co-operation can be successful only by exchanging electronic documents and using the change representation functions and/or the commenting function", all 26 participants were asked to choose one of the five alternatives (Strongly disagree, Disagree, Neutral, Agree, Strongly agree). The result says that half

of them (10 for agree, and 3 for strongly agree) agreed upon it (See Figure 5). It is somewhat surprising because it is well-known to be hard to produce a document together without face-to-face or voice communication. Therefore, we think that one should not undermine the potential of co-authoring mainly through electronic tools, the commenting and change representation functions. Rather, we might have to say that its potential is widely open particularly in the situations where physical meetings are difficult to hold, or when we need to reduce those meetings.

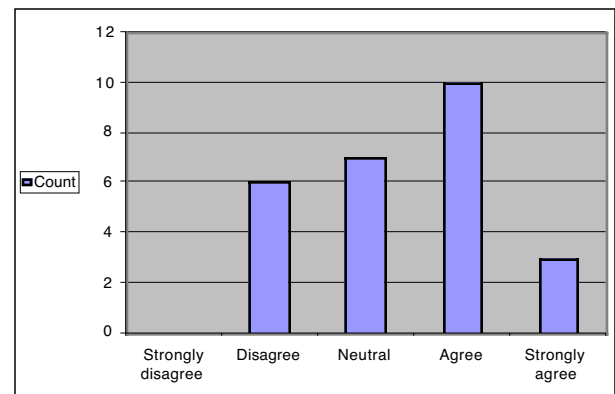


Figure 5. Perceived possibility of co-authoring without face-to-face meetings.

3.6.6 Summary

Reviewers in the RI model spent more time than in the traditional model. However, we did not find any difference of patterns of discussion at the meetings according to the amount of discussion at four different levels mentioned previously. Reviewers' understanding of text was almost the same between the two conditions. Writers' sense of help from reviewers with respect to the improvement of text showed no difference. On the basis of the results, we believe that the aspects like understanding text and sense of help are not so much associated with the technology. Rather, social interaction and partners' motivations affect those aspects more. As an advantage of it, nevertheless, many subjects believe that the RI on-line collaboration can be successfully carried out only by using the change representation function and the commenting function, when it is difficult to have physical meetings.

4. CONCLUSIONS

This paper has introduced the RI on-line reviewing model, and presented its effects according to multiple dimensions such as patterns of revisions and comments, use of paper documents, emotion, cognitive loads, and understanding and improvement of text.

Reviewers' direct revision of the original on-line text usually concerns small details in it, while they discussed parts of documents demanding higher level revisions by other ways of communicating such as face-to-face meetings and the commenting function. This result in fact provokes an interesting issue in the design: Can new ways of representing changes, if possible, promote higher level revisions, or do the system designers take it for granted that the reviewer mainly tends to make revisions at a low level when working under the RI setting? As a matter of fact, the answer to the question is strongly related to understanding of what are the main reasons that such a tendency of revising is observed. For example, is it because reviewers are likely to avoid making high level revisions regardless of ways of presenting changes? Is it, otherwise, because current change representation tools influence the tendency? While recognizing the limits of this study in presenting evidence-supported reasons regarding the tendency, we hope that more research on it will be carried out.

In the experiment, eleven of all 13 reviewers have used the commenting function together with the change representation function. Because the commenting function was just an optional choice of the reviewers participated, we believe this result at least reveals a certain strong connection between change representation and comment. As previously mentioned, reviewers have used the commenting function when there was a difficulty of formulating exact words. Therefore it is of great importance to consider how to design change representation tools in connection with commenting functions.

The study also showed that the RI reviewing with the change representation technology did not affect the improvement of text quality as much as we anticipated. The degrees of writers' sense of improvement of the text by the reviewer's assistance, and of reviewers themselves' understanding of text in the RI setting, were not significantly different from those of the traditional way. We thus conclude that who reviews the text is more important than what technology is used for reviewing, while we believe change representation tools are certainly useful to treat details at a low level in the text. Further, change representation tools were not a

source of emotional conflicts in terms of sense of feeling hurt or being hurt.

The subjects involved in the study, the first year university students, have less experience of writing than professionals who spend much time writing. They might be also less engaged in the given task, since their purpose to participate in the lab assignment was mainly to get the grade for the course. Therefore, it is interesting to investigate the cases where writing professionals are involved, in order to see how they use electronic reviewing tools and how the results in the study match those cases.

Although this study concerns the RI reviewing model only, a further exploration of the WI model is also requested to understand its effects and access the differences between the WI and the RI models, and between the WI on-line reviewing and the traditional reviewing on paper.

Finally, on-line reviewing is a growing practice. However, reviewing with pen and paper is still extensive. Instead of predicting what medium will prevail in the future, we believe that it is more desirable to understand the differences between the two media and how they are harmonized and used during the reviewing process. Though we have shown that paper copies of documents were rarely used while reviewing in the RI mode, we feel that more new research should be opened up, which deals with the issues of how each medium is used while reviewing.

REFERENCES

- Ede, L. and Lundford, A. (1990): *Singular Texts/Plural Authors: Perspectives on Collaborative Writing*. Southern Illinois University Press.
- Gould, J. D., Alfaro, L., Barnes, V., Finn, R., Grischowsky, N., Minuto, A., Salaun, J. (1987): Reading is slower from CRT display than from paper: Attempts to isolate a single variable explanation. *Human Factors*, 29, 269-299.
- Kim, H. C. and Severinson Eklundh, K. (1998): How academics co-ordinate their documentation work and communicate about reviewing in collaborative writing, Technical report TRITA-NA-P9815, NADA, August 1998.
- Kim, H. C. and Severinson Eklundh, K. (2000): Change representation in collaborative writing, Technical report TRITA-NA-P0005, NADA, March 2000.

Lay, M. M. and Karis, W. M. (1991): *Collaborative Writing in Industry: Investigations in Theory and Practice*. Baywood Publishing Company, Inc.

Neuwirth, C. M., Chandhok, R., Kaufer, D. S., Erion, P., Morris, J. and Miller, D. (1992): Flexible Diff-ing in a Collaborative Writing System. In *proceedings of CSCW '92*, 147-154.

O'Hara, K. and Sellen, A. (1997) A comparison of reading paper and on-line documents. In *proceedings of CHI 97*, 335-342.

Rada, R. (1996): *Groupware and Authoring*. (Eds.) Academic Press Inc.

Santos, A. (1995): *Multimedia and Groupware for Editing*. Berlin Heidelberg: Springer-Verlag.

Severinson Eklundh, K. (1992): Problems in achieving a global perspective of the text in computer-based writing. *Instructional Science*. 21, pp. 73-84.

Sharples, M. (1993): *Computer Supported Collaborative Writing*. (Eds.) London: Springer-Verlag.

ACKNOWLEDGEMENTS

Special thanks to Stina Hållsten and Per-Anders Forstorp who made it possible to organize and conduct the laboratory study, and to Professor Kerstin Severinson Eklundh who reviewed this paper with many helpful comments.