

# Pasting and Encoding: Note-taking in Online Courses

Aaron Bauer, Kenneth Koedinger  
*Human-Computer Interaction Institute*  
*Carnegie Mellon University*  
{abauer, koedinger}@cmu.edu

## Abstract

*Studies have shown that both the act of note-taking and the use of notes for review can promote learning. Many note-taking applications have been developed for computer-based learning content. In general, they include advanced annotation functionality, and are geared toward supporting collaboration and discussion. Though these devices have been shown to change note-taking behavior, their effect on learning has not been evaluated. The goal of our research is to evaluate the effect of specific features of note-taking applications on behavior and learning, in order to develop guidelines for advanced note-taking applications that promote learning. These applications could be used as the basis for a variety of educational activities, including collaboration. In this paper, we present the results of an experiment evaluating a basic feature of note-taking technology: copy-paste. Our findings indicate that copy-paste functionality can be detrimental to learning. We describe potential implications of these results for the developers of note-taking applications.*

## 1. Introduction

Note-taking is a common and important practice for students both in the classroom and while reading text [11,17]. Empirical evidence exists for two potential benefits of taking notes. Process benefits refer to learning gains obtained from the act of taking a note. Review benefits accrue from the ability students have to review their notes. Kiewra found that approximately half of process studies showed positive process effects, while three-fourths of review studies found positive review effects [6]. While these studies deal mostly with note-taking during lecture, they have been replicated for note-taking while reading (for example, [7]).

Technology allows new methods of both delivering content and supporting note-taking. Most researchers

have focused on building multi-featured annotation applications. They are most often studied in the context of discussions and collaborative work. The devices themselves, as well as the collaborative contexts in which they are used, have been shown to change the process of note-taking. Unfortunately, these changes in behavior, as well as their impact on learning, have not been studied in detail.

There is a need for foundational research investigating the relationship between digital note-taking functionality, behavior and learning. This research will allow us to develop applications that both promote learning and can serve as the basis for collaborative tasks. In this paper we present the results of an experiment indicating that inclusion of one basic feature, copy-paste, appears to reduce long-term retention of material. We describe the potential implications this has for the design of more advanced note-taking applications.

## 2. Note-Taking Technology

Note-taking devices built for computer-based content, such as online courses, have generally supported full-featured annotation. For example, Wang and colleagues developed an application that allowed students to annotate a document, add comments to the annotation, and classify the annotation [15]. This application was used as the basis for student discussion of the learning content, and to promote mentoring [16]. Marshall and Brush evaluated a similar application aimed at supporting student discussion of digital reading materials [9].

The focus on using note-taking as the basis of digital collaboration is pervasive. Kurhila et. al. evaluated a device that made annotations part of the public record, which students used to produce group reports [8]. Others have developed joint note-taking applications for lectures. In the Livenotes project, groups of

students work together to record one set of notes during lecture. [5]

These devices have been shown to affect how students take notes. For example, students in one study on annotating while reading found it difficult to make general summary comments when they were limited to attaching comments to highlighted areas. They also changed their note-taking style when they knew they would share their annotations with other students. [9]. Developers have successfully curtailed specific note-taking behavior based on the features included in their devices (verbatim note-taking, in the case of [13]).

We do not know how these changes in behavior affect learning. Though some evaluations exist, they either lack statistical significance (see [5]), or are not true experiments, comparing different populations and years (see [8]). Regardless, evaluating multi-featured applications would be difficult, as one could not ascribe learning gains to individual features, some of which may not even be related to note-taking.

### 3. Our project

The goal of our research is to understand how specific note-taking features affect both behavior and learning. Therefore, our project iteratively evaluates individual features. Our initial tool behaves as a simple text-editor [figure 1]. Students using it can record ideas, but cannot draw pictures, annotate course materials (highlighting or underlining), or anchor notes in the course material (marginal notes).

One previous study found that, when compared with note-taking using pencil-and-paper, a similar tool produced superior review benefits [12]. In our own preliminary study, we did not find any difference on an immediate multiple-choice post-test. However, we did

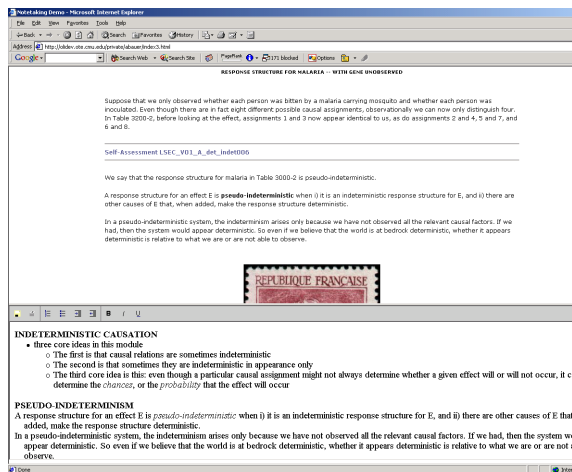


Figure 1: The Note-taking application

find that students took advantage of the copy-paste feature to produce notes with far more words, of greater verbatim proportion, than did students who used pencil-and-paper [1].

Many researchers regard verbatim note-taking as a negative behavior. Though verbatim notes can be quite common [3], many associate it with shallow processing. Few studies have investigated the impact of these notes on learning outcomes, though one [2] did show students who were forced to take notes verbatim showed poor retention. However, as students report strategies for verbatim note-taking [14], there is a question of external validity with regards to this result.

In our previous study, the use of copy-paste also produced more wordy notes. Copy-paste decreases the cost, with regards to both time and effort, associated with wordier notes. Therefore, students using this feature use more words to record an idea than students using paper. Some studies find wordy notes to be detrimental to retention [4], while others find wordier notes to increase retention [10].

The copy-paste feature appears to be an important function of the tool. In addition, the process of copy-paste is similar to popular annotation functionality, in that the interaction technique is selection, rather than typing. We thus designed a study to evaluate the impact of the copy-paste feature. This study compares notes taken using pencil-and-paper with notes taken using a tool that either permits or prohibits copy-paste. Students are tested at three intervals: immediate, delayed, and delayed with review. We believed that prohibiting pasting would result in fewer overall notes, of a lower verbatim proportion. We also believed that while pasting would reduce retention scores due to inhibited processing, the accuracy of pasted notes would result in better performance at review.

### 4. Study Method

A total of 69 subjects from several local universities were recruited by means of a posting to a subject-recruitment website. SAT Math scores were found to account for a significant amount of variability in test scores, and were used in our learning analyses. Unfortunately, 17 participants did not report their scores, so we were only able to include the data of 52 subjects in our analysis.

The study took place on two days separated by one week. On the first day, students reviewed a module in an online course in Causal and Statistical Reasoning, taking notes using one of the three note-taking conditions mentioned above: Paper, Tool-noPaste, and Tool-Paste. They were then administered an immediate

post-test. A week later, they returned and were given two additional tests, allowing five minutes for review between tests.

We expanded upon the 12-item multiple choice test from the previous study. The new tests contained a total of 28 questions, approximately two-thirds of which were multiple choice items. The remaining third were free-response items. As we did not have data to match the test-items statistically, we counterbalanced their presentation.

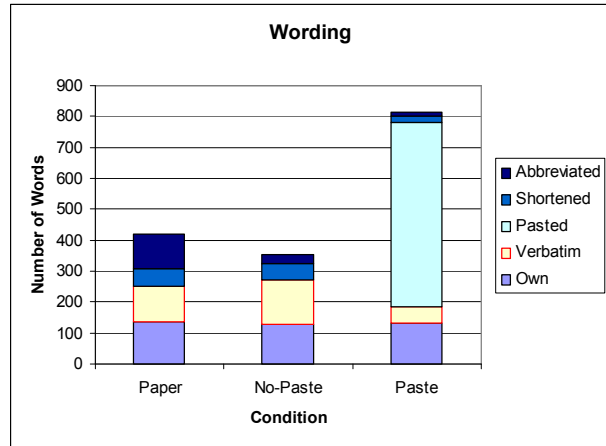
Participants' notes were broken down into ideas, and then were coded with regards to wording. Four wording categories were developed: "Verbatim" notes were of the exact form as in the module. "Abbreviated" notes had the same words in the same order, but could include abbreviated words or leave out conjunctions, such as "and", or simple prepositions, such as "to." "Shortened" notes had the same words in the same order, but could leave out major words or sections of one to five words. "Own" notes either used completely different words or word orders. Key ideas, or those identified as critical by the course developer, and which formed the basis of the tests, were also coded.

A limitation of this research is that one researcher coded all the data. We believe this limitation is mitigated by the fact that notes were entered into a spreadsheet, and later coded blind to the condition that produced the notes. Future work will address this issue.

## 5. Results

ANOVAs were performed on all measures of notes quantity, with condition as the only independent variable. We found significant differences with regards to both note-quantity and wording. With regards to overall note quantity, the Paste condition produced more words than the other two conditions [ $F(2,49)=7.2$ ,  $p=.001$ ] (figure 2), however it only produced significantly more ideas [ $F(2,49)=3.9$ ,  $p=.02$ ] than the no-paste condition (figure 3). The conditions were only marginally different, in the same direction, with regards to the number of key ideas recorded [ $F(2,49)=2.64$ ,  $p=.08$ ]. The difference between words and ideas is indicative of a brevity difference. Pasters produced by far the most wordy notes [ $F(2,49)=11.48$ ,  $p<.0001$ ].

With regards to wording, Paste was once again characterized by a large amount of verbatim notes. It produced significantly more verbatim words [ $F(2,49)=13.6$ ,  $p<.0001$ ] and ideas [ $F(2,49)=3.3$ ,  $p=.04$ ] than the other conditions. While there was no significant difference with regards to number of words in the own words category [ $F(2,49)=.03$ ,  $p=.96$ ], paper produced more own ideas than the other two conditions



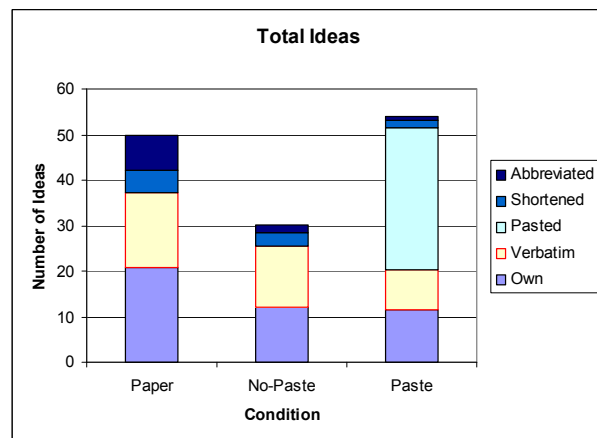
**Figure 2: Words by Condition**

[ $F(2,49)=3.9$ ,  $p=.02$ ]. Though there were some effects of abbreviated and shortened notes, they composed a small enough portion of notes that they could not be analyzed with regards to learning.

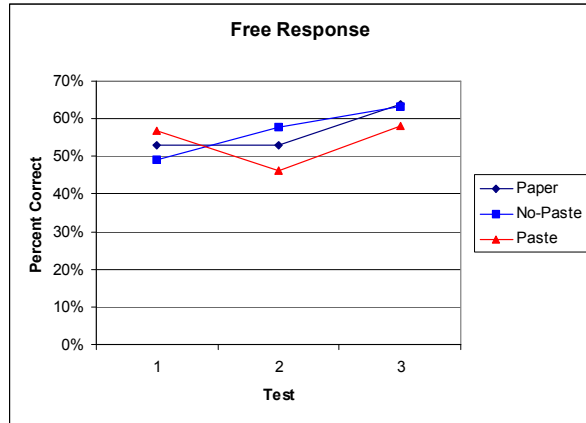
We conducted ANOVAs for all individual tests, including condition, test form and SAT-Math in the model. The latter was included because it was found to account for a large amount of the variability in learning results. No significant effects were found on any of the individual multiple choice or free response tests.

We also performed repeated measures analyses on tests one and two, looking for retention (or process) effects, and on the second and third test, looking for review effects. The overall review effect was marginal for multiple choice tests [ $F(1,34)=2.8$ ,  $p=.09$ ], and significant for free response tests [ $F(1,34)=15.7$ ,  $p=.0004$ ]. However, review appears to be a robust effect, as we found no condition by test interactions.

We did not find an overall process effect for either test type. However, there was a significant test by condition interaction for free response questions



**Figure 3: Total Ideas by Condition**



**Figure 7: Free Response scores. The overall interaction is marginal ( $p=.09$ ), however the forgetting interaction between the first two tests is significant.**

[ $F(2,34)=4.1$ ,  $p=.02$ ]. Students using the paste condition showed poorer long term retention, forgetting significantly more information than did the other conditions.

We see two major factors distinguishing Paste from the other conditions that could explain poor learning results. Pasters took more verbatim notes, and their notes were significantly more wordy than those of the other conditions. As described above, verbatim note-taking and wordier notes could help explain reduced retention. As each free response question is tied to a specific key idea, we can treat each question as a single data point, associating it with a specific wording and brevity. We could then create a statistical model with Condition, SAT-Mean, Test (including the first two), and Brevity/Wording included as independent effects, and controlling for Subject as a random effect (as each subject would be associated with multiple questions).

We only had enough data to evaluate the “own” and “verbatim” wording categories. We did not find a significant wording by test interactions on the first two tests [ $F(1, 376)=1.4$ ,  $p=.23$ ], indicating that the wording of key ideas was not responsible for retention loss. Brevity produced somewhat more interesting results. We expected that more wordy notes would be associated with reduced retention, as the pasters had produced wordier notes and forgotten more. However, we found a marginally significant effect in the opposite direction [ $F(1, 876)=3.2$ ,  $p=.07$ ]. It turns out that Paste is the only condition that does not follow this trend. Neither No-Paste nor Paper showed any retention losses between for wordy items. However, Paste did show these retention losses. It appears that using more words to express ideas did not derive the same

encoding benefits for the Paste condition that it did for the other conditions.

## 6. Discussion and Design Implications

We found that copy-paste functionality substantially changes behavior. Participants took advantage of the copy-paste facility to produce a larger proportion of verbatim notes, which were much wordier than notes in the other conditions. This did not appear to affect review benefits obtained from notes, as review was robust across conditions. However, it did appear to affect process benefits, as participants in the copy-paste condition forgot more free-response items between the first and second test. This appears to be due to the wordiness of the notes. Students in the paste condition did not receive the same retention benefits from wordy notes as did students in the other conditions.

This is an important result for designers of note-taking applications. Not only do the features included change the way students take notes, they have the potential to impact learning. Careful consideration should be taken when designing the feature-set. The specific result here has direct implications for features that share a similar interaction technique to copy-paste. Annotation tools, in particular, are used to select text in a fashion similar to that of copy-paste. When used to take notes, they may suffer similar learning effects. However if we want to design tools to promote encoding, it is important to understand why copy-paste is detrimental to encoding. We have two hypotheses, which when tested will allow us to develop guidelines for annotation tools.

The first step in note-taking is identifying an item to record. This is identical regardless of the note-taking condition. In the second step, the note-taker must decide which parts of the item to transcribe. This is where we see the first difference between note-taking conditions. The pasters’ tendency to transcribe wordier notes meant they could record all words associated with an idea, no matter how inconsequential. In fact, it is possible that students identified key ideas by surface structure alone, not even reading the material. In the other conditions, the costs associated with increased wordiness led students to record fewer words. This may be an indication of increased attention to critical details, which in turn increases the processing of the idea. This is the *attention hypotheses*: copy-paste reduces attention to critical details of ideas, thus reducing processing.

After selecting what will be recorded, the note-taker must transcribe the idea. In the paste-condition, this requires two simple motions, which take a limited

amount of time. In the other conditions, students actually have to generate their notes by typing or handwriting. They may receive some processing benefit from the generation of their own notes. This is the *generation hypothesis*: the act of generating the words in one's notes increases processing benefits.

These two hypotheses have different implications for tools that use selection as the main interaction technique. For example, if the attention hypothesis is correct, students should not be allowed to select large segments of text; instead they should be forced to identify key segments for annotation. If the generation hypothesis is correct, it may be appropriate to require students to add information to their annotations, perhaps through comments, in order to make them more active note-takers.

## 7. Future Work

We are currently exploring the above hypotheses in order to allow us to develop more specific guidelines for note-taking applications. We will include two new interfaces in our experiments. The first will limit the amount of text students can select, in order to require students to attend to the critical details of the ideas. The second allows students to select from a set of options for each note they take, in order to test the generation hypothesis. We will also conduct more naturalistic experiments in actual online courses, in order to evaluate the impact of tool features in realistic settings.

The results of these experiments will be used to produce an annotation note-taking tool that supports learning. This tool can form the basis for the same collaborative tasks as in the tools described above.

## Acknowledgements

This research was conducted under the Open Learning Initiative at Carnegie Mellon University, which is funded by a grant from the William and Flora Hewlett Foundation. The authors would like to thank Richard Scheines, John Rinderle, and Bill Jerome for their support of this work.

## References

[1] Bauer, A., and Koedinger, K. "Developing a Note Taking Tool from the Ground Up". *Ed-Media 2005*. *AACE Press*, 4181-4186.

- [2] Bretzing, B. H., & Kulhavy, R. W. "Note-taking and depth of processing." *Contemporary Educational Psychology* 4, (1979) 145-153.
- [3] Bretzing, B.H., Kulhavy, R.W. "Note-taking and Passage Style", *Journal of Educational Psychology* 73, 2 (1981) 242-250
- [4] Howe, M.J. "Using students' notes to examine the role of the individual learner in acquiring meaningful subject matter." *Journal of Educational Research* 64, 2 (1970), 61-63
- [5] Kam, M., Wang, J., Iles, A., Tse, E., Chiu, J., Glaser, D., Tarshish, O., and Canny, J.F. "Livenotes: a system for cooperative and augmented note-taking in lectures." *Proc. CHI ACM Press* (2005), 531-540
- [6] Kiewra, K.A. and DuBois, N.F. "Note-taking Functions and Techniques." *Journal of Educational Psychology* 83, 2 (1991) 240-245
- [7] Kulhavy, R.W, Dyer, J.W., & Silver, L. "The effects of note taking and test expectancy on the learning of text material." *Journal of Educational Research*, 1975, 68, 522-528
- [8] Kurhila, J.; Miettinen, M.; Nokelainen, P.; Tirri, H.; "The role of the learning platform in student-centred e-learning" *IEEE ICALT'04*, p.540 – 544
- [9] Marshall, C.C. & Bernheim Brush, A.J. "Exploring the Relationship between Personal and Public Annotations." *DL 2004, ACM Press* (2004), 349-357
- [10] Nye, P., Crooks, T., Powley, M., & Tripp, G. "Student note-taking related to university examination performance." *Higher Education* 13, (1984). 85-97.
- [11] Palmatier, R.A., Bennet, J.M. "Note-taking habits of college students." *Journal of Reading* 18, (1974) 215-218
- [12] Quade, A. M. An assessment of retention and depth of processing associated with notetaking using traditional paper and pencil and on-line notepad during computer-delivered instruction. *AECT'95* (1995)
- [13] Truong, K. N., Abowd, G.D., and Brotherton, J.A. "Personalizing the Capture of Public Experiences." *Proc. UIST'99, ACM Press* (1999), 121-130
- [14] Van Meter, P.M., Yokoi, L., Pressley, M. College "Students' Theory of Note-taking Derived From Their Perceptions of Note-taking." *Journal of Educational Psychology* 86, 3 (1994) 323-338
- [15] Wang, C.Y., Wei, F.H., Chao, P.Y., Chen, G.D. "Extending e-books with contextual knowledge recommenders by analyzing personal portfolio and annotation to help learners solve problems in time." *IEE ICALT'04* p.306 – 310
- [16] Wei, F.H., Lee, L.Y., Chen, G.D. "Supporting adaptive mentor by student preference within context of problem-solving learning" *IEEE ICALT'04* p.789 – 791
- [17] Williams, R.L., Eggert, A.C. "Note-taking in college classes: Student patterns and instructional strategies." *The Journal of General Education* 51, 3 (2002), 173-199