

# Remembrance of Things Tagged: How Tagging Affects Human Information Processing

Raluca Budiu, Peter Pirolli, Lichan Hong

Palo Alto Research Center

3333 Coyote Hill Road, Palo Alto, CA 94304 USA

{raluca.budiu, pirolli, hong}@parc.com

## ABSTRACT

We describe an experiment investigating how tagging affects human memory. We compare a low-cost tagging by mouse-clicking interface (Click2Tag) to traditional tagging by typing and to a baseline, no tagging condition. Our results suggest that tagging is beneficial for memory. However, the two tagging methods impact distinct aspects of memory: tagging by typing promotes elaboration of read contents and, thus, enhances free recall, whereas tagging by clicking strengthens the memory traces by repeated readings of relevant words in the text and, thus, improves recognition of facts from the text.

## Author Keywords

Tagging, memory, interaction cost, social bookmarking

## ACM Classification Keywords

H5.m. Information interfaces and presentation: Miscellaneous. H5.3 Group and organization interfaces: Web-based interaction.

## INTRODUCTION

In recent years, there has been an explosion of social bookmarking systems (e.g., del.icio.us, diigo, ma.gnolia). These systems allow users to generate labeled links (*tags*) to content encountered on the Web; the social aspects of these systems emerge from implicit or explicit sharing of tags. As evidence of this growth, the Pew Internet Survey [10] reports that 28% of U.S. Internet users have tagged or categorized content such as videos, photos, news stories, or blogs. Theoretical explanations [4,8], discussed below, attribute the growth of participation in socially mediated systems partly to the reduction in the costs of producing sharable content (e.g., bookmarks and tags). Consequently, in the development of our own social tagging system (SparTag.us [7]) we have been motivated to develop

techniques that lower the costs of producing tags and other annotations. However, theories of memory [1] also suggest that changes in the kind and amount of tagging effort involved may affect how well people remember the original content or tags. We want to avoid tagging techniques that provide lower costs of interaction but at high costs of remembering the material.

In this paper, we unravel how different techniques for producing tags to Web content affect individual memory. We are also generally interested in whether the act of tagging increases memory for content. In particular, we performed an experimental contrast of the (lower interaction cost) Click2Tag technique, developed for SparTag.us, against a standard (higher interaction cost) type-to-tag technique, similar to ones used in popular tagging systems such as del.icio.us. These both were contrasted with a baseline condition of no tagging. Click2Tag allows users to simply mouse-click words in a text to have the words become tags for the content. Type-to-tag allows users to type their own tags for the content. We examined the effect of these techniques on both recognition and recall tests for the original material. As we discuss below, previous memory research [1,5,9] leads us to hypothesize that Click2Tag would produce better recognition of the original material, but type-to-tag would produce superior recall. These results are predicted by differences in the way the techniques either strengthen or elaborate the memory traces for the original content. We also discuss the implications of this trade-off for the functionality and design of social tagging systems.

## TAGGING COSTS

### Tagging Cost and Participation

Social information foraging theories [8], as well as microeconomics theories of networked information economies [4] predict that as the costs of production of shareable knowledge (e.g., tags) are driven down, more individuals will participate and reap greater net benefits. Thus, reductions in the cost of tagging will improve the value of the system to the individual user. More tags, and presumably more useful tags for the individual, will be generated as more people join a social tagging site.

### Click2Tag: Lowering the Cost of Tagging

To lower the cost of tag production, SparTag.us makes each content word on a Web page clickable. As users click words, the words are added to a tag list. The assumption is that this technique (further called Click2Tag) provides a low-cost method for tagging in comparison to the more standard type-to-tag technique. In type-to-tag, users can generate any string as a tag and enter it into a tag list. The difference is that this involves the additional costs of mentally generating the tag plus the cost of typing.

### Cognitive Costs on Memory

One worry about Click2Tag might be that it trades off costs in tagging time for cognitive costs to subsequent memory. Human memory research suggests that Click2Tag vs type-to-tag may have different effects on subsequent memory.

The *strength* of a memory trace is a theoretical construct that captures the relationship of practice to memory performance [1]. Repeated practice increases strength, and strength decays as a function of time since last practice. Reaction times and accuracy on recognition tests both improve with strength (e.g., [9]). Click2Tag appears to encourage users to re-study the original content and, thus, increase the strength of memory traces for that content. Indeed, eye tracking data from pilot studies in our lab suggest that users often read a passage and then re-scan the passage to seek out words to click. In comparison to simply reading text (with no tagging required), we expect Click2Tag to produce better recognition performance.

*Elaborativeness* of encoding of material also generally improves aspects of memory. If people are asked (or provided) with additional information that is highly semantically related to the content they are studying, they typically show superior recall to content that has been processed less elaborately [5]. The memory traces that elaborate the original content provide additional retrieval routes to recall the content. This is because self-generated elaborations have some high likelihood of being re-generated at recall time as a retrieval route to the content. Since type-to-tag requires users to generate tags to associate with the original content, we expect it to produce more elaborative encodings and to improve recall performance.

## EXPERIMENT

### Method

**Participants.** We recruited 20 employees of our company to participate in this study. The participants were compensated with \$20 Amazon gift certificates.

**Materials.** We selected 24 passages from news articles as well as from various web pages on the Internet. The passages reflected a variety of topics (medicine, education, general science, aviation, history, etc). On average, the passages were 267 words long (ranging from 253 to 279).

**Procedure.** Participants went through three study-recall blocks. A study-recall block had two parts: in the first part,

participants performed 6 study trials, and in the second part, participants performed 6 memory trials.

*Study trials.* In each study trial, participants read a passage, selected randomly from the list of 24 passages. Participants were instructed to read at their own pace, but if they spent more than 2 minutes on a trial, they were automatically moved to the next trial. The trial could belong to one of three conditions as follows:

- **No-tags:** In this condition, no tagging was performed.
- **Click2Tag:** Participants had to tag the passage with relevant words by clicking on words from the passage. The tags were displayed in a box under the passage and could not be modified by the participants.
- **Type-to-tag:** Participants had to tag the passage with any relevant tags that they could generate, and type those tags in a box under the passage.

*Memory trials.* After 6 study trials, the participants completed 6 memory trials. A memory trial had two components, presented in the following order:

- **Recall:** Participants were given two cues about one of the 6 passages they had previously studied (e.g., “One of the passages you read was about Christmas and Santa Claus”). Then they had to remember and type as many facts from that passage as possible. There was a time limit of 1.5 minutes per passage for this phase.
- **Recognition:** Participants had to answer 6 true/false sentences (3 true and 3 false) about the passage they had just recalled. They had 1 minute to answer all questions. The order of presentation of questions was randomized. The order of passages within a block was randomly generated for each participant, as was the assignment of a passage to a particular block or condition. The passages in the memory trials were presented in a random order, unrelated to the order in which the passages were studied. The only constraint was that there be 6 passages per block, and within each block, there be 2 passages per condition. At the end of each block, subjects were given feedback about their overall performance so far in the recognition tests. We did not give feedback after individual recognition trials.

For each trial, we measured the study time, the number of tags, the number of facts recalled and the recall time, the recognition time per question and the recognition accuracy.

### Results

We performed ANOVAs with subjects as the random factor using Block (0,1,2) and Condition (no-tags, Click2Tag, type-to-tag). When there was a correlation with study time, we performed ANCOVAs with study time as a covariate.

### Encoding

Table 1 shows the average study time per condition and per block. There was a significant effect of block on the study time ( $F(2,38)=8.74$ ,  $p<0.001$ ), with participants taking longer to study the passages as they went through the experiment. There was also a significant effect of condition

	Block 0	Block 1	Block 2	Average
No-tags	68.43	83.85	88.29	<b>80.19</b>
Click2Tag	92.24	96.81	95.97	<b>95.01</b>
Type-to-tag	94.06	100.46	101.30	<b>98.61</b>

Table 1. Study times (seconds).

( $F(2,38)=45.08$ ,  $p<0.001$ ) and a significant block-by-condition interaction ( $F(4,76)=3.91$ ,  $p<0.01$ ). Contrasts showed that participants spent less time in the no-tags condition than in the Click2Tag condition ( $p<0.001$ ) and in the type-to-tag condition ( $p<0.001$ ). These results pointed to a time cost associated with the tagging conditions, but that cost was manifest mostly in the beginning of the experiment. In the later blocks, the cost difference washed away, possibly indicating that participants decided to spend extra time on studying the text in the no-tags condition.

People tended to attach more tags ( $p<0.001$ ) in the Click2Tag condition (6.33 on average) than in the type-to-tag condition (4.08 on average), suggesting that they made use of the ease to tag in the Click2Tag condition to attach more tags faster than in the type-to-tag condition.

### Recall

**Time to recall.** There was no effect of the time to recall – people took the same amount of time to recall the content of the passages in all conditions (77.66 seconds on average).

**Number of recalled facts.** The left panel in Figure 1 shows the number of recalled facts. Due to an error in the experiment software, we did not have access to any recalled materials for 202 trials (out of 360 trials). We had two raters rate the number of different facts per passage in a recall trial. The interrater correlation was 0.87. The number of recalled facts was negatively correlated with the study time ( $-0.35$ ,  $p<0.001$ ), so we performed an ANCOVA with study time as a covariate. We obtained a significant effect of block ( $F(2,15)=4.10$ ,  $p<0.05$ ) and a marginally significant interaction between block and condition ( $F(4,30)=2.42$ ,  $p<0.07$ ). Contrasts showed that, in Block 0, participants recalled more facts in the no-tags condition than in both the Click2Tag condition ( $p<0.05$ ) and the type-to-tag condition ( $p=0.07$ ). In Block 2, participants in the tagging conditions (Click2Tag and type-to-tag combined) recalled more facts ( $p<0.05$ , one-tail  $t$ -test) than in the no-tags condition. Moreover, participants in the type-to-tag condition recalled more than in the other two conditions ( $p<0.05$ , one-tail  $t$ -test). In the middle block, the contrast between tagging and no-tags is also significant ( $p<0.05$ , one-tail).

These results show that there was an overall learning effect

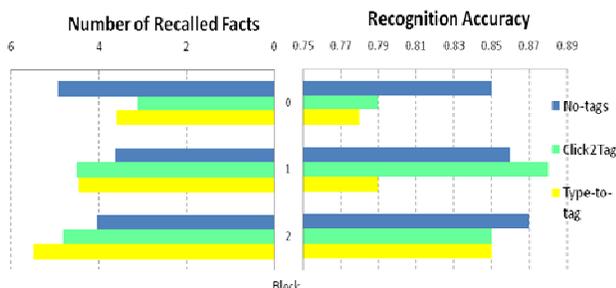


Figure 1. Recall and recognition.

through the experiment, with people getting better at recalling as the experiment went on. They also indicate that, whereas people started by recalling better the passages that they did not tag (presumably because they used their time to study as opposed to tagging), at the end of the experiment, they achieved best performance in the type-to-tag case. Moreover, both tagging conditions were better than no tagging by Block 1. This supports the hypothesis that, in the type-to-tag condition, participants produced meaningful elaborations that helped them at recall.

### Recognition

We initially ran AN(C)OVAs with the truth of the sentence being judged as an additional factor. However, truth was at best marginally significant, and always in the expected direction (i.e., faster/more accurate true judgments than false judgments). So, in the interest of space, we only report the AN(C)OVAs with Condition and Block as factors.

**Accuracy.** The right panel in Figure 1 shows the recognition accuracies. The correlation between the study time and the accuracy at recognition is low ( $-0.02$ ) and not significant, and the ANCOVA with study time as a covariate preserves the results, therefore we only report the ANOVA. We found a significant effect of condition ( $F(2,33)=5.79$ ,  $p<0.01$ ), and a significant interaction between condition and block ( $F(4,74)=2.34$ ,  $p<0.05$ ). People responded correctly on 81% of the trials in the type-to-tag condition; this was significantly lower than the 86% accuracy in the no-tags condition ( $p<0.01$ ) and the 85% accuracy in the Click2Tag condition ( $p=0.05$ ). When we examined the source of the interaction, we obtained that no-tags was better than type-to-tag ( $p<0.05$ ) in Block 0 and in Block 1 ( $p=0.05$ ). Click2Tag was better than type-to-tag in Block 1 ( $p<0.01$ ). None of the differences were significant in Block 2, suggesting that there may be a learning effect in using type-to-tag effectively for the task in this study.

These results show that Click2Tag did not impair memory, in spite of reducing the cognitive costs of processing. People recognized facts equally well in the no-tags and Click2Tag condition. However, they were slightly less accurate when they typed (although they improved over time). This supports our hypothesis that Click2Tag strengthens facts in memory, and thus has better effects on recognition than type-to-tag. As in the case of recall, no-tags was better in the beginning than type-to-tag because presumably people did not need to pay the extra cognitive cost of thinking of relevant words, and they devoted their effort to studying the passages.

**Recognition times.** Table 2 shows the recognition times. There was a significant correlation ( $0.34$ ,  $p<0.001$ ) between study time and recognition time, so we report results from an ANCOVA with study time as a covariate.

The condition was significant ( $F(2,32)=3.29$ ,  $p<0.05$ ), as was the condition by block interaction ( $F(4,73)=3.66$ ,  $p<0.01$ ). Contrasts showed that Click2Tag led to faster

	Block 0	Block 1	Block 2	Average
No-tags	4.81	5.06	4.85	4.91
Click2Tag	4.97	4.43	4.65	4.66
Type-to-tag	5.07	5.13	4.89	5.03

**Table 2. Recognition times (seconds).**

recognition times than both no-tags ( $p=0.07$ ) and type-to-tag ( $p < 0.01$ ). However, if we ran separate tests for each block, this difference was valid only in Block 1 (difference with no-tags:  $p < 0.01$ ; with type-to-tag:  $p < 0.005$ ).

Since Click2Tag was faster in Block 1 than the other conditions and also more accurate than type-to-tag and as accurate as no-tags, there was no time-accuracy trade-off involved. The lower time of Click2Tag validates our assumption that Click2Tag is a low cost interaction.

## DISCUSSION

Overall, our results show that tagging helps memory, although in different ways for recall and for recognition. In the case of recall, type-to-tag is most beneficial: at the end of the experiment people used type-to-tag efficiently to form elaborations that helped them with recall. Tagging conditions were better than no tagging for recall, with type-to-tag being most helpful. At recognition, Click2Tag was more effective than typing initially, presumably because participants who clicked fixated more on relevant words in the text, thus strengthening their memory traces. No-tags surpassed type-to-tag at recognition accuracy in the first two experimental blocks, but in the end type-to-tag was no different from the other conditions, indicating that participants got better at doing the task over time.

Our results are similar in spirit to note-taking research. One study [3] compared note-taking using typing versus pasting and found that pasting led to better learning than typing. However, [2] found that note-taking using pasting combined with typing made participants record more ideas, but also forget more. Other research [6] has shown that verbatim notes (presumably similar to Click2Tag) led to poorer learning than paraphrased or summarized notes.

## CONCLUSIONS

We started this paper with an argument over the importance of reducing the interaction cost for social tagging systems. However, in reducing the interaction costs, we must ensure that we do not also increase cognitive costs. We showed that Click2Tag, a method that reduces the tagging cost by allowing users to click words in order to tag content, does not impair memory performance. In fact, it can be beneficial to recognition memory, due to strengthening of relevant words in text. In contrast, typing tags encourages text elaborations and improves recall of tagged content. However, it leads to more recognition errors, compared to a no-tag condition or to Click2Tag. Type-to-tag is a top-down procedure. It induces users to fit the content into their own "ontology" and retain only those facts matching their view of the world. On the other hand, Click2Tag is more bottom-up, content-driven. People tag with relevant words in the passage, paying less attention to their own ontology. One

question that remains to be explored is whether the two techniques will lead to different tag "folksonomies". Another related question refers to the impact of the tagging technique on information retrieval. If Click2Tag leads to a proliferation of tags, would those tags contain less information and impact the ease of retrieval? How about the combination of Click2Tag and typing – how would it compare to typing or clicking?

We showed that tagging helps memory. From a designer's perspective, Click2Tag and type-to-tag may serve different purposes, as they affect memory in different ways. For systems that promote learning, it is possible that Click2Tag may offer faster, more content-driven tagging, which can be recognized better at low cost. For systems that emphasize personal organization of material, type-to-tag may help users' subsequent memory by incorporating their own interpretations and ontologies.

## ACKNOWLEDGMENTS

We thank Michelle Gumbrecht for useful comments; Leila Takayama for revising the materials; our participants.

## REFERENCES

1. Anderson, J. R. (2000). *Learning and Memory: An Integrated Approach*. John Wiley, New York, NY.
2. Bauer, A. & Koedinger, K. (2006). Pasting and Encoding: Note-Taking in Online Courses. *Proc. ICALT'06*, 789-793.
3. Bauer, A. & Koedinger, K. (2007). Selection-Based Note-Taking Applications. *Proc. CHI'07*, 981-990.
4. Benkler, Y. (2002). Coase's Penguin, or Linux and the Nature of the Firm. *Yale Law Journal*, 112, 367-445.
5. Bradshaw, G. & Anderson, J. (1982). Elaborative Encoding as an Explanation of Levels of Processing. *J. of Verbal Learning and Verbal Behavior*, 21, 165-174.
6. Bretzing, B. & Kulhavy, R. (1979). Note-Taking and Depth of Processing. *Contemporary Educational Psychology*, 4, 145-153.
7. Hong, L., Chi, E., Budiu, R., Pirolli, P., & Nelson, L. (2007). SparTag.us: Low-Cost Webpage Tagging at the Paragraph Level. *Under review*.
8. Pirolli, P. (2007). *Information Foraging Theory: Adaptive Interaction with Information*. Oxford University Press, New York, NY.
9. Pirolli, P. & Anderson, J. (1985). The Role of Practice in Fact Retrieval. *J. of Experimental Psychology: Learning, Memory, and Cognition*, 11, 136-153.
10. Rainie, L. (2007). 28% of Online Americans Have Used the Internet to Tag Content. *Pew Internet & American Life Project*. [http://www.pewinternet.org/pdfs/PIP\\_Tagging.pdf](http://www.pewinternet.org/pdfs/PIP_Tagging.pdf). Retrieved Sept, 2007.