HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments ISSN: 1795-6889

www.humantechnology.jyu.fi Volume 6 (2), November 2010, 229–248

TEMPTING TO TAG: AN EXPERIMENTAL COMPARISON OF FOUR TAGGING INPUT MECHANISMS

Mark Melenhorst
Novay
Enschede, the Netherlands

Lex van Velsen
University of Twente
Department of Technical and Professional Communication
Enschede, the Netherlands

Abstract: *Tagging helps achieve improved indexing and recommendation of resources (e.g., videos or pictures) in large data collections. In order to reap the benefits of tagging, people must be persuaded to label the resources they consume. This paper reports on a study in which four different tagging input mechanisms and their effect on users' motivation to tag were compared. The mechanisms consisted of a standard tag input box, a chatbot-like environment, a bookmarking mechanism, and a "tag and vote" game. The results of our experiment show that the use of the nonstandard tagging input mechanisms does not affect users' motivation to tag. In some instances tagging mechanisms were found to distract users from their primary task: consuming resources. Persuading people to tag might be accomplished more effectively by using other motivating tagging mechanisms (e.g., tagging games), or motivation could be created by explaining the usefulness of tagging.*

Keywords: *information retrieval, tagging, motivation, interaction design, experimental research.*

INTRODUCTION

One of the core innovations in Web 2.0 applications has been the possibility for users to share content, such as Web references, photos, videos, or presentations, with other users. When users upload their content, they can supply it with self-devised keywords. Depending on your privacy preferences, these applications allow other users to retrieve your content by means of these keywords. Asking readers to describe their content in their own words seems to be a sensible thing to do, since the inclusion of self-devised keywords, also known as *tags*, are beneficial for the indexing of the resource. Tags describe a resource (e.g., a scientific paper, photo, or video) in layman's terms or add information to the content (Berendt & Hanser, 2007).

Tagged resources are easier to find in large collections (Melenhorst & van Setten, 2007). Furthermore, the addition of tags makes it easier for other users of a system to understand the

URN: NBN:fi:jyu-201011173093

contents of the resource quickly, since tags provided by peers have been found to be more descriptive of the content than keywords provided by external professionals (Matusiak, 2006). And finally, tags reflect the interests of a user, upon which the personalization of output can be based. In this context, tags are seen as indicators of users' characteristics and contexts. An overview of how tags can be employed for user models is provided in Wartena (2010).

Of course, tagging also has its limitations (see Mathes, 2004). Tags can be ambiguous, which can make searching for a resource difficult (e.g., searching for pictures of *spoon* can result in both tableware and the rock band Spoon), and a lack of synonym control can result in a set of almost identical tags describing a resource (e.g., a picture of the Alps can be tagged as "mountain" and "mountains"). The many successful on-line services that provide the option to tag resources have made it clear, however, that the limitations do not have to stand in the way of the advantages.

A look at the popular Websites on which one can tag shows that, more often than not, platforms that facilitate the sharing of nonprofessional content do not force the users to tag, despite the advantages we just described. In the cases where the use of tags is supported, the choice of whether or not to tag is the user's. This does not have to prevent people from tagging. Several on-line services, like del.icio.us and Flickr, acquire many tags even though the choice to tag or not is left to the user. By not obligating users to tag the resources they are consuming or uploading, services such as these avoid the risk of irritating users by forcing them to devise tags, at the cost of less user-generated metadata. However, new services, or services with a small number of users, do not have such a well-developed folksonomy (a large set of tags that describes resources) to their avail, which hinders advanced indexing. In cases like these, the service will want to create a folksonomy with a small number of people or in a relatively short amount of time. One way to achieve this goal is to tempt people to tag.

Several studies have delved into the issue of how to motivate people to voluntarily tag content so as to increase the amount and diversity of metadata. The study described in this paper contributes to the discussion on how to tempt users to tag on-line content by comparing four different tagging input mechanisms. The results of this comparison can inform designers regarding the usefulness of creating and implementing tagging input mechanisms that are different from the standard input field currently used in most applications that facilitate the use of tags.

The remainder of this paper is organized as follows. In the next section, we will discuss the relevant literature. This is followed by the presentation of the experimental setup and the four tagging input mechanisms that were compared. We then discuss the results of the comparison, and finalize this paper with our conclusions.

THEORETICAL BACKGROUND

Tagging is a way to add metadata to a resource. These resources can be very diverse, ranging from videos (e.g., www.youtube.com), to songs (e.g., www.last.fm), to books (e.g., www.librarything.com). The various definitions of social tags emphasize the freedom of users in assigning the keywords they deem fit. Wu, Zubair, and Maly (2006) define tags as freestyle descriptors of a resource, while Sen et al. (2006) define them as short free-form labels used to describe items in a domain. The literature concerning tagging shows two main

streams: evolution and effect studies, and studies delving into the matter of how to motivate people to tag on a voluntary basis.

Evolution, Effect, and Motivation

Evolution studies try to map the development of a collection of tags, associated with a large information collection (e.g., Golder & Huberman, 2006). With this knowledge, it is possible to predict how large collections of tags come about and evolve. As a result, one can determine at what point in their evolution tag collections can be used effectively to improve other functionalities, such as searching. Effect studies on tagging, meanwhile, have shown that the inclusion of tags in content searched via a search engine increases the quality of search results (Melenhorst, Grootveld, & van Setten, 2008; Morrison, 2008). However, it is not correct to say that an increase in the number of tags provided will result in an increase of quality of search results. According to Chi and Mytkowicz (2008), this quality is dependent on the diversity of the tags the users provide. Therefore, an application that is aimed at tempting users to tag should focus not only on quantity, but also on diversity.

Because tagging is an activity performed most often by a small minority of a user population (Marlow, Naaman, Boyd, & Davis, 2006), another strand of research is focused on motivating all users to tag. Heckner, Heilemann, and Wolff (2009) divide the motivations for tagging into two main categories: personal information management and resource sharing. While the former is intended to efficiently and effectively store and retrieve content, the latter class of motivations is more social: the key point is to distribute valuable information with other users. A generic but more fine-grained overview of user motivations to tag can be found in Marlow et al. (2006).

Other researchers have focused on system-specific motivations to tag, assuming that generic motivations may not comply with a system-specific context. Several studies have focused, for example, on users' motivations to tag photos (Ames & Naaman, 2007; Nov, Naaman, & Ye, 2008). Interestingly, the large majority of user motivations are instrumental. In other words, people do not tag because they think it is fun, but they hope to achieve a different goal by means of tagging. For example, on MovieLens,[1] users primarily tag movies to create an overview of the kind of movies they have seen (Sen et al., 2006).

Early motivation research and literature has focused on general human behavior. The seminal work by Maslow (1943), for example, describes how human behavior is the result of needs. These needs range from the physiological (e.g., breathing or food), to love (e.g., friendship) to self-actualization (e.g., creativity). However, human motivation cannot be explained simply by a set of motivational factors, but rather is context dependent (Ryan & Deci, 2000). Situational motivation can be explained by means of the notions of intrinsic and extrinsic motivation. One is intrinsically motivated when someone does something for fun or out of interest, while one is extrinsically motivated when the executed behavior is instrumental to another goal (Deci & Ryan, 1985).

Deci & Ryan (1985) argue that intrinsic and extrinsic motivations are interrelated when rewards are used to stimulate external motivation. When rewards affect the perceived self-determination and perceived competence, they may also increase intrinsic motivation. If not, rewards potentially threaten intrinsic motivation. There is considerable evidence for the negative effect of tangible rewards (e.g., money) on intrinsic motivation (Deci, Koestner & Ryan, 1999).

Various on-line activities are initiated because of the different intrinsic or extrinsic motivations. Using the Internet in the context of work is primarily the result of extrinsic motivation: It should be useful for achieving a certain goal (Teo, Lim, & Lai, 1999). On-line shopping, on the other hand, is an activity that is fun for many of people, and hence, is intrinsically motivated for them (Shang, Chen, & Shen, 2005). Finally, the use of medical Websites has been found to be both intrinsically and extrinsically motivated (Logan et al., 2000).

Tagging Applications: Designing for Motivation

Several publications discuss research that explores different techniques for increasing users' motivation to tag various kinds of context. Drenner, Shen and Terveen (2008) conducted a study with users of MovieLens in four conditions. In one condition, users were not forced to tag. These users were in turn assigned to one of two groups: those who did and those who did not receive a screen in which they could tag on a voluntary basis. In the other two conditions, users had to tag either 5 or 25 movies. More people did not complete the more intensive condition (tag 25 movies) than the less intensive conditions. However, the new users who were forced to tag more movies were more fanatic taggers during their subsequent use of MovieLens, even if forcing the new users to tag did not influence the quality of tags. This study shows that it is possible to shape future tagging behavior of users by means of a specific interaction design.

Other researchers have focused on positively influencing tagging behavior by focusing system design upon the needs and wishes of the target user group. The mobile photo application Zonetag (Ames & Naaman, 2007), for example, gives users the possibility to skip the option to tag their photos. It can also provide users with tag recommendations. The recommendations are based upon the tags of other photos that have been taken at the same location and tags that the user has submitted in the previous 24 hours. These recommendations were found to increase the number and diversity of tags. Similar promising findings for tag recommenders have been found in the context of recommending tags for blog posts (Sood, Owsley, Hammond, & Birnbaum, 2007) and in social bookmarking systems (Jäschke, Marinho, Hotho, Schmidt-Thieme, & Stumme, 2008).

An interesting application that approaches tagging as a game is the ESP game (Von Ahn & Dabbish, 2004). In this game, two persons are shown the same photo that they tag simultaneously without seeing the other person's tags. When an identical tag is submitted by both persons, they are awarded points. Since its launch, the ESP game has proven to be a success, and collected more than 10 million tags in the first few months (Von Ahn, 2006). The success of this game has shown that presenting tagging input mechanisms as a form of entertainment has the potential of greatly improving the number of tags provided by users.

Social Communities: Interface and Interaction Design

In the literature, several implications for the design of motivating interface and interaction design for social communities can be found. Tagging is often done in the context of a social community. Therefore, tagging can be regarded as a voluntary contribution to such a community. YouTube is, for example, more than a huge database of movies, but also a place for people to gather and socialize (Lange, 2008).

A study concerning how to motivate social community members to rate movies identified that members are more willing to rate a movie when the interface shows that the community will benefit from this rating. This effect was even larger when the interface indicated that members themselves would benefit from their actions (Rashid et al., 2006). Another study showed that movie community members are more willing to invest time and effort when they are given specific goals (e.g., rate 16 movies in the next week; Beenen et al., 2004). In the case of a social peer-to-peer downloading service, it has been found that rewarding active members with an upgrade in their membership status (e.g., silver or gold members) and an improvement in the service, such as increased privileges, motivates users to be more active in the community (Cheng & Vassileva, 2005).

The works mentioned in this section have generated some useful guidelines for the design of social communities that can also be used in the context of applications that facilitate the possibility to tag (see, e.g., Preece, 2000). However, most guidelines concern general interaction design. In order to improve the design toolkit for tagging applications, we investigated how different tagging input mechanisms affect users' motivation to tag.

INPUT MECHANISMS SELECTION AND EXPERIMENT SETUP

This study is the next in a line of work that focuses on user motivation to tag videos. In this section we will first briefly discuss our previous studies. Then we will present the four different tagging input mechanisms that we compared. We will conclude this section with describing the experimental procedure we applied.

Previous Work: Video Tagging and Motivation

In the first stage of our work into user motivations to tag video content, we focused on putting together a list with users' possible motivations to tag a video on the Internet. Based on focus groups, we compiled a list with possible motivations related to indexing, socializing, and communicating (Van Velsen & Melenhorst, 2008).

Next, a large group of intensive Internet users ranked these possible motivations for two cases: uploading a video onto an on-line news Website and watching a video on an on-line music community (Van Velsen & Melenhorst, 2009). In both cases the motivations related to indexing were the main motives to tag an uploaded or watched movie. The motivation "tagging as a means to make others able to find a movie" was in both cases the most important motivation of all. Interestingly, affinity with the subject at hand did not lead to a higher motivation to tag: People tag certain video content to achieve another goal (e.g., improved indexing of a movie) not because they think a video is funny or interesting. Based on these findings, one can say that video tagging by means of a traditional tag entry box is extrinsically, rather than intrinsically, motivated.

The next step in our research was to take these insights, translate them into tagging input mechanisms, and to put these to the test. These tagging input mechanisms were the result of several brainstorming sessions.

Brainstorming Sessions

As a first step in developing the different tagging mechanisms to be compared in our experiment, two brainstorming sessions were held. The first session was held with a class of 25 third-year college students majoring in digital communication at the Hogeschool Utrecht (a university of applied science in the Netherlands). First, it was explained what tagging entailed. Next, groups of five to six students were assigned to discuss and come up with ideas for motivational tagging systems. To promote the elaboration of the ideas, the ideas from one group were passed to another group after which all of the ideas were further discussed and new ideas were generated in the group as a whole.

The second brainstorming session was held with 11 people: six experts from the fields of digital communication, cross media studies, and usability, who were teachers at the aforementioned school, and five student researchers in the field of digital communication. The process was the same as the process that was followed in the first brainstorming session.

The results of these brainstorm sessions was a long list of ideas. These ideas, listed in Table 1, represent potential means to motivate users to contribute tags.

Several ideas that served a purpose other than motivating users to contribute tags were left out. From the extensive list of ideas, three ideas were selected and further elaborated into working prototypes, hereafter referred to as tagging mechanisms. *Bookmarking* was selected because of our earlier research: We found that personal indexing was the most important motivation for users to tag (Van Velsen & Melenhorst, 2009). Therefore, it could be considered the most promising mechanism.

The *chatbot/chatbox* was selected because of its attempt to transform tagging into chatting. This is an activity in which many Internet users engage because of a social, intrinsic motivation (Stafford, Stafford, & Schkade, 2004).

The *tagging game* was selected because it appeals to users' motivation for competition and play (Marlow et al., 2006). In addition, voting for tags may improve their quality.

Descriptions of Tagging Mechanisms

The selected mechanisms were integrated into a Web environment specifically designed for this study. In this environment, the outline of the study, the experimental environment, and the concept of tagging were explained to the user. After this introduction, the user interacted with the interfaces one by one. For each mechanism, the user was asked to watch two videos, presented in a YouTube-like style. For each mechanism, help information was made available and, if necessary, the researcher could assist the participants. In Appendix A, a screen dump is displayed for each of the mechanisms.

Condition 1: Tag Box

Rationale. This mechanism does not have a specific motivational quality. It represents the way tagging is implemented in most Websites today. As such, it is the baseline against which the other mechanisms in this study are compared.

Table 1. Long List of Tagging Input Mechanisms.

Idea	Description	Intended Motivation
Bookmarking	*Tags could serve as input for a bookmarking system. By tagging certain content, they would be able to find it again more easily. The system would automatically order and display the "tagged" favorites by type of content, such as videos about pets or videos containing spoken language. As such, this idea resembles Del.icio.us.*	Future retrieval, contribution and sharing
Personal homepage	Introducing new videos on personal home pages of social network sites.	Self-presentation
Involving the social network	On the Website, one could give users the ability to create a personal friends list, or allow users to put themselves on an uploader's friends list. When the uploader shares a new video, an e-mail would be sent to his or her friends containing a link to the video in question and the request to create some tags for it, or to comment on the resource. Such a subscription method is already being used on YouTube.	Attract attention
Reward system	When users assign tags, they could be rewarded with more (related) content.	*
Commercial tagging	Following some review Websites, financial rewards could be given, for instance, based on a share of advertisement revenues	*
Chatbot and chatbox	*Users could be invited to chat about a video. When no other users are watching a video at the same time, a chatbot invites users to talk about the video. Tags can be derived from the chatlogs.*	(see section on Condition 2)
Tagging game 1	*Users could tag and subsequently vote for tags that they think are good. Votes are counted and prominently displayed. As such, users are encouraged to compete with each other to generate many high-quality tags.*	Competition and play self-presentation
Tagging game 2	Two players could simultaneously see the same image and try to come up with the same tags. If they do, they would be awarded points. The high scores of individual players would be displayed on the site prominently. Such a game is already present in the form of the ESP game.	Competition and play self-presentation
Tagging game 3	"Where is Waldo" is a game in which a little figure is hidden in the to-be-tagged resource. Multiple quick frames of Waldo could be hidden in a video. After the video is complete, the user can indicate at which frames Waldo appeared or, rather, what happened when Waldo appeared. If more people give the same answer (e.g., tags), they receive points, and their description of the scene becomes a tag.	Competition and play self-presentation
Tagging game 4	When key frames are extracted from the videos, they can be compared against Flickr photos. A game could ask users to identify the differences between the Flickr photo and the YouTube clip, from which tags can be extracted.	Competition and play self-presentation
Tagging game 5	After a video has ended, the system could present the user with a small quiz. For example, when there is an image of a cat walking across the street, a quiz question could be: "What did the cat pass on his way to the other side?" From these answers, tags could be derived.	Competition and play Self-presentation

Note. Selected ideas in italic. Motivations indicated by * give rewards in cash or in content, which are not covered by taxonomies for tagging incentives. Intended motivations are based on Chi & Mytkowicz (2008) and Marlow et al. (2006).

Functionality. This mechanism consists of the usual text box with a Tag button. No specific attempts were made to encourage users to tag here.

Condition 2: Chatbot

Rationale. Apart from the suggested mechanisms, the brainstorm sessions led to the conclusion that the propensity to tag could increase when tagging as an uninspiring activity is avoided. Earlier research (e.g., Sen et al., 2006; Van Velsen & Melenhorst, 2009) has shown that there is no intrinsic motivation for tagging, but that it is only done to achieve a certain objective. The chatbot idea does not encompass a classical tagging activity but replaces it with something that could appeal to an intrinsic motivation: to get involved with other people and friends. When chat functionality is offered next to a movie clip, it can be assumed that conversations revolve around this movie clip. Tags can be derived from the chat protocols by extracting the most salient and often-used words. Statistical techniques can be used to filter out off-topic conversations (e.g., Wartena & Brussee, 2008) and to distinguish topic related words from other salient terms (Wartena, 2010).

Functionality. Users can chat about the video in a chat window that is presented next to the video clip. When no other users are on-line, users can chat with a chatbot (an artificial-intelligence-based computer that can communicate with users more or less like a human being) that invites the users to tell him what the video clip is about. However, this was presented as an invitation in order to avoid pressuring the users to use the chatbot.

Condition 3: Bookmarking

Rationale. In a previous study (Van Velsen & Melenhorst, 2009), we found that personal indexing or indexing for others are the most important motivations for users to engage in video tagging. This prototype draws on this motivation. To a certain extent, it resembles Del.icio.us.

Functionality. Users can organize their bookmarks into folders and tag them. Subsequently, they can retrieve their bookmarks via these tags. Thus, in addition to a basic tagging mechanism, it allows users to organize their content by means of tags.

Condition 4: Tag & Vote

Rationale. This mechanism was created on the assumption that people like it when they can display their competence by being named in a high score list.

Functionality. Users can tag video clips and rate other users' tags by voting for what they think is the best tag. Tags receiving more than three votes are visible to other users. Users are able to see how many votes their tags received and what their position in a high score list is.

Experimental Set-Up

We constructed an experiment in which we evaluated the motivational effect and the appreciation of the interfaces with the implemented mechanisms that were described in the previous section.

Participants

Forty participants were informally recruited. They were, on average, 23.4 years old (*SD* = 5.0): 29 were male and 11 were female. They were all college students. However, students attending programs in digital communication, information science, and related disciplines were not allowed to participate in the study, since their prior knowledge about the topic may have interfered with the objectives of the study.

All but one of the participants use the Internet on a daily basis. Typical Web 2.0 applications are not used regularly, apart from YouTube and Hyves (a Dutch Facebook-like community). Twenty-five of the 40 participants used YouTube once a week or more, while Hyves is used once a week or more by 26 out of 40 participants. No one used Del.icio.us, one participant used Flickr once a week or more, and only six used Last.fm once a week or more.

With respect to their on-line activities, the results show that only three participants tag more than once a week, while 29 participants never tag. Sixteen out of 40 participants contribute to a forum once a week or more, while instant messaging is most popular: 28 of 40 use IM messaging more than once a week.

In sum, for this group of participants, popular social tagging applications are used only to a small extent, indicating that tagging is not so widespread among the group of participants, who may be considered as frontrunners with regard to the use of Web 2.0 applications. This result is consistent with our earlier work (Van Velsen & Melenhorst, 2009), in which we found that only 20% of the information elite knew what tagging was about.

Procedure and Tasks

The experimental procedure was completed one person at a time and consisted of the steps listed below. The entire procedure was presented within the electronic environment. The language used in the interfaces was Dutch, even though some of the movie clips were in English. Even though this environment guided the participant through the experiment, a researcher was available for questions and technical assistance.

1. *Introduction*. The experiment's steps were explained to the participant. Two things were assessed here: the participant's study subject and his/her familiarity with tagging.
2. *Reading an introduction to tagging*. Next, the core concepts and principles for tagging were explained. Each participant had to read this introduction, even if the user was already familiar with tagging: The purpose was to create a common understanding of tagging.
3. *Experimenting with the mechanisms and watching the video clips*. The participant went through all four prototypes. The order in which the prototypes were presented was randomized. For each prototype, two video clips were shown. After each video clip, a short survey was administered with questions concerning the participant's appreciation of the video and his/her propensity to tag the video clip. Following the second video in each condition, the participant was questioned additionally about the appreciation of the tagging mechanism in question and about the added value of tagging when presented this way.

4. *Survey*. The final part of the experiment consisted of a survey with questions regarding demographics and use of Web 2.0 applications.

Materials

Eight short YouTube video clips were selected and paired for each condition: Four clips were meant to entertain users and four clips were of an informative nature. They all lasted about three minutes. The titles and URL's of the videos can be found in Appendix B. The clips were presented by means of YouTube's embedded player within the ePaxperimental environment.

Data Collection

Using surveys, we collected the following data by means of short surveys:

1. Appreciation for the content, using 5-point scales and a holistic mark on a scale from one to ten (after each clip)
2. Propensity to tag (after each clip and after each mechanism): the participant's inclination to tag using the mechanism provided
3. Perceived usefulness and usability of the tagging input mechanisms (after each mechanism)
4. Background characteristics (at the end of the study)

The surveys are displayed in Appendix C.

RESULTS

In this section we discuss the results of our study. First, we address the results regarding the tagging input mechanisms. Then, we address the role of the content and its influence on the propensity to tag.

Appreciation of Tagging Input Mechanisms

After the users watched the two video clips per tagging mechanism, they were asked to provide a generic evaluation of the mechanism. We first tested whether the appreciation for the different mechanisms differed. The results are shown in Table 2. The bottom row represents the test-value of the within-subjects effect resulting from a repeated measures analysis with "tagging mechanism" as within-subjects factor.

As Table 2 shows, the scores regarding usefulness items received moderate scores. The usability items were more positively scored with means around four. Contrary to our expectations, the propensity to tag is above the neutral point of 3. We think this is somewhat surprising since the literature suggests that a small percentage of Internet users engage in tagging. Hence, we expected values to be lower than the neutral point. The added value of tagging the movie clips is considered relatively low, with a score slightly below the neutral point of 3.

Table 2. Effect of Bookmarking Mechanism on Perceived Usefulness and Usability.

Condition	Usefulness			Usability			
	Tag propensity	***Added value***	***Use in real life***	***Ease of use***	***Learna-bility***	***Compre-hension***	***Fun to use***
Control condition [C]	3.1 (1.1)	2.5 (.9)	3.0 (1.1)	4.1 (.7) ^{B}	4.2 (.6) ^{B}	4.1 (.8) $^{Ch, T, B}$	3.0 (1.2) ^{T}
Chatting [Ch]	3.5 (1.1)	2.9 (1.1)	3.2 (1.1)	4.1 (.8) ^{B}	4.2 (.8) ^{B}	3.7 (1.1)	3.5 (1.0)
Tagging & Voting [T]	3.2 (1.1)	2.6 (1.1)	3.3 (1.1)	3.8 (.9)	4.0 (.8)	3.5 (1.0)	3.6 (.9)
Bookmarking [B]	3.3 (1.1)	2.7 (1.1)	3.3 (1.1)	3.5 (1.0) ^{C,Ch}	3.6 (1.1) $^{C, Ch}$	3.2 (1.1)	3.4 (1.0)
F-Value[a]	1.34	1.98	.81	5.18 **	5.35 **	7.27 ***	3.28 *

Note. Values for the prototype-evaluations could range from 1 to 5. Standard deviations between parentheses. Significant differences between one mechanism and another are indicated by a superscript that refers to the first character(s) of the other mechanism. The significance level is .05.
[a] statistical significance: * = . at .05 level; ** at .01 level; *** at. 001 level

With regard to the perceived added-value of tagging, no statistically significant differences were found between the tagging mechanisms. The control condition (with a basic tag box) did not result in a lower perceived added value in comparison to the other tagging mechanisms.

Table 2 does show some differences in the perceived usability of the input mechanisms. The bookmarking mechanism was less easy to use and had a more troublesome learnability than the control condition and the chatbox. Not surprisingly, the control condition was the easiest to understand. The "fun to use" criterion did yield somewhat ambiguous results. Significant differences between the control condition and the tag and vote condition were found, but not between the control condition and the other conditions. This is somewhat surprising since we expected all mechanisms to be more fun to use than the control condition. In the case of the chatbot, this effect may have been caused by the absence of other users to chat with: Chatting with other users will probably be more appreciated than chatting with an automatic chatbot.

To get a better understanding of the relationship between propensity to tag and usability, we computed correlations between ease of use, learnability, instant comprehension, and the propensity to tag. In the bookmarking condition, each of the usability criteria was positively correlated with the propensity to tag ($.39 < r < .57$; $p < .05$). For the voting condition, learnability was positively correlated with the propensity to tag ($r = .37$; $p < .05$). For the chatbox condition and the control condition no correlations were found. These results suggest that usability can affect users' intention to tag.

Appreciation of Movie Content

The tagging input mechanisms cannot be considered in isolation from the content they are presented with since the content may influence users' appreciation of the mechanisms. Therefore, we investigated the relations between the content and the input mechanisms. After each video clip, the appreciation of the video clip was assessed by means of six items, derived

from Norris & Colman (1994). Participants had to award up to 5 points on each of the 6 appreciation items. Cronbach's alpha for the scale was .93. Table 3 displays the scale means.

We performed a MANOVA analysis with the tagging input mechanism as an independent variable, and average content appreciation and propensity to tag as dependent variables. Familiarity with the movie clip, presentation order of the tagging mechanisms, the type of movie clip, and the position of the subject in the experiment were introduced into the model as covariates. The model proved to be statistically significant, $F(2, 75) = 191.99$, $p < .001$.

Further inspection of the between-subjects results showed that the tagging mechanism had a statistically significant effect on the appreciation of the content, $F(3, 75) = 5.64$, $p < .01$. However, as Table 3 shows, advanced tagging mechanisms do not lead to a higher appreciation for the content than the simple tag box: The differences between the control condition and the other mechanisms were not significant.

Furthermore, the video clips were appreciated less in the bookmarking condition (Bonferroni post-hoc test; $p < .01$) and the voting condition (Bonferroni post-hoc test; $p<.01$), compared to the chatting condition, but not in comparison with the control condition. The lower appreciation for tagging & voting and bookmarking could be the result of distraction, since the items assessing usability pointed out that the participants found the bookmarking and the voting mechanism more difficult to understand than the mechanism in the control and the chatbot condition. This could have interrupted their attention to the video clips, possibly affecting their appreciation for the content. In contrast to the ratings, the propensity to tag was not affected by the tagging mechanism, $F(3, 75) = 2.50$, n.s. In other words, each of the mechanisms resulted in the same propensity to tag.

To further explore the relationship between the propensity to tag and the appreciation for the content, we computed correlations between both variables. There proved to be a significant correlation between the appreciation for the content and the propensity to tag ($r = .32$, $p < .001$). The next step was to construct a regression model with propensity to tag as the dependent variable, and appreciation for the content as the independent variable. The model proved to be significant with a R^2 of .10, $F(1, 318) = 36.04$, $p < .001$) with a highly significant B ($B = .38$; $t = 6.00$, $p < .001$). Introducing tagging mechanism as a second independent variable led to a nonsignificant Beta ($B = -.-6$, $t = -1.10$, n.s.).

Table 3. Content Appreciation and Propensity to Tag by Individual Clip Type and Condition.

Tagging mechanism	**Content appreciation**			**Propensity to tag**		
	Clip Type			***Clip Type***		
	I*	E	Avg.	I	E	Avg.
Control [C]	3.4 (1.0)	3.6 (.9)	3.5 (.9)	2.7 (1.3)	2.5 (1.3)	2.6 (1.3)
Chatting [Ch]	3.8 (.9)	3.6 (.9)	3.7 (.9)	2.7 (1.1)	2,7 (1.2)	2.7 (1.2)
Bookmarking [B]	3.3 (1.1)	3.0 (1.2)	$3.1(1.1)^{Ch}$	2.3 (1.1)	2.0 (1.2)	2.2 (1.2)
Tagging and Voting [T]	3.3 (1.0)	2.9 (.9)	$3.1(1.0)^{Ch}$	2.3 (1.3)	2.3 (1.2)	2.3 (1.2)
Average	3.4 (1.0)	3.2 (1.0)	3.3 (1.0)	2.5 (1.2)	2.4 (1.2)	2.4 (1.2)

Note. Significant differences between one mechanism and another are indicated by a superscript that refers to the first character(s) of the other mechanism.
The significance level is .05. * I stands for Informational video; E stands for entertaining video.

CONCLUSIONS

This paper presented an experiment in which we compared four different tagging input mechanisms and investigated how each mechanism affected users' motivation to tag. The four mechanisms were the result of two brainstorming sessions with students of digital communication and Web 2.0 professionals. The mechanisms consisted of a control condition (a standard tagging text input box with a tag button), a chat window in which one can chat with other users or a chatbot and from which tags are derived automatically, a del.icio.us-like bookmark mechanism and, finally, a mechanism by which one could tag and then vote for "good" tags that were kept on a high score list. The experimental results show that the different input mechanisms tested in the experiment do not lead to different perceptions of the added value of tagging, nor do they affect the users' propensity to tag. The appreciation of the content to tag was affected by the tagging input mechanism. This might have been the result of presenting a relatively complicated tagging input mechanism, which might have distracted the participants from the video content.

Our results indicate that implementing "fancy" tagging input mechanisms that utilized a chatbot, a voting mechanism for the best tag, or a bookmarking feature do not lead to a higher motivation to tag. Of course, this finding does not rule out more advanced tagging input mechanisms in general. It is possible that a different tagging design could have been perceived as prettier or more interesting by the participants which, on its turn, might have influenced usability or motivation scores. In the Evolution, Effect and Motivation section, we discussed several other promising tagging input mechanisms (like the ESP game). The implementation of these tagging input mechanisms might well be more fruitful than the mechanisms tested in this study and might motivate users to tag more than by use of a standard tag entry box. However, our opinion is that the prototypes used in this study are a proper and realistic embodiment of the ideas behind them. Therefore, although a different prototype might have marginally influenced results, we think that the general trend that can be observed in our results holds. Of course, a definitive verdict on this issue can only be realized by means of a replication study using our prototypes and new prototypes that represent the same idea.

It is possible that the tagging input mechanisms we tested did not affect user motivation because users need to be convinced of the added value of tagging in a different way. The point at which it is explained to users what tagging is and what purposes it can serve could be a crucial moment. After reading or watching this explanation, users will have to decide for themselves whether they find tagging worth the effort. Only after making this decision, then, should the users be confronted with the tagging input mechanism. The various mechanisms are used to achieve the benefits of tagging and are instruments for that, but they are not motivators. In other words, taggers do not make the decision to tag or not on the basis of the tagging input mechanism presented to them. Because tagging is done with a higher goal in mind, it is not as strongly affected by the interface and interaction design of the tagging input mechanism as we presumed it to be.

The best moment in time at which users can be persuaded to tag and in which form depends heavily on the users' contexts and goals. These contexts and goals can be identified by applying a user-centered design approach. In such an approach, (potential) users are consulted as early as possible in the design process, after which their characteristics, wishes,

and contexts lead the design (Gould & Lewis, 1985). The functional design and the interface and interaction design that are the fruit of this design approach may well be highly motivating. Therefore, case studies on user-centered design of motivating tagging interfaces will be a very welcome addition to the tagging literature.

The data provide us with some evidence that the mechanisms that were considered more difficult to understand might have disturbed the participants' appreciation for the movie clips. However, the setup of the study does not allow for a thorough analysis of the relation between these factors. Because we think that the interplay between usability and motivation to tag is an important determinant of tagging motivation, future research should delve deeper into this relationship.

However, we can safely conclude that when designing tagging mechanisms, software developers have to be careful to pay attention to usability criteria as well as their motivational quality. In any case, tagging input mechanisms should not be too intrusive, as also noted previously by Sen et al. (2006).

Interestingly, the results of the study suggest that appreciation of the content was positively co-related to a users' propensity to tag. This result is contradictory to the results we found previously (Van Velsen & Melenhorst, 2009), where higher affinity with the content to be tagged did not lead to a higher propensity to tag. We can only conclude that the relation between these two factors is unclear at the moment. Further research will have to shed more light on this relation.

IMPLICATIONS FOR PRACTITIONERS

In this study, several tagging input mechanisms were compared to test which one encouraged users to supply most tags in the context of video tagging. These involved a chatbot mechanism (chatting with a chatbot so tags can be derived from the logs), a bookmarking condition (where users can organize their bookmarks by using tags), and a tag & vote condition (where users can tag videos and rate other users' tags), as well as a traditional tagging method (tag input box) as a comparison condition. The results show that the advanced tagging input mechanisms do not improve users' motivations to tag. Therefore, designers of tagging applications have two options. The first is to use a standard tag input box where users can type tags. This is an easy and cheap solution and yielded the same results as the more advanced tagging input mechanisms we tested. The second option is to motivate users to tag by means of implementing other tagging input mechanisms. For instance, tag recommenders or tagging as a game as incorporated in an ESP game[2] have been shown to be more promising than the mechanisms we tested in this study.

Taking into account the results from our earlier studies, reported in Van Velsen and Melenhorst (2009), we can state that motivating users' to tag video content requires the careful selection of the right focus and instruments. The primary motivation to tag a video is to make a video (or other medium) easier to find for others or yourself. This activity can be simplified by using the use of tag recommenders (Melenhorst et al., 2008). An alternative approach may be to make tagging fun by means of a game. Which approach works best will be dependent on the system, the context, and the user. Even though the interplay between these dependencies is a research topic in itself, system developers need to determine the best

approach by exploring the system's context of use and identifying its target group. Based on this knowledge the most appropriate approach can be selected.

A second finding from this study is that usability is related to users' motivation to tag. Therefore, it is important to ensure that a tagging input mechanism is usable. In order to achieve this goal, we recommend applying a user-centered design perspective while creating tagging input mechanisms and testing the interface and interaction design of a mechanism before launch.

Finally, in the case of video tagging, it is very important that tagging input mechanisms do not distract the user too much from watching the video. A novel or relatively complex tagging input mechanism might prove too distracting and has implications for the tagging process and effectiveness. It is paramount to present the users' primary goal—that is, watching a video—as the main activity in the interface and interaction design of a video application.

ENDNOTES

1. See www.movielens.org
2. See www.espgame.org

REFERENCES

Ames, M., & Naaman, M. (2007, April-May). *Why we tag: Motivations for annotation in mobile and online media*. Paper presented at the Computer-Human Interaction conference (CHI 2007), San Jose, CA, USA.

Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P., & Kraut, R. E. (2004, November). *Using social psychology to motivate contributions to online communities*. Paper presented at the Computer Supported Cooperative Work conference (CSCW '04), Chicago, IL, USA.

Berendt, B., & Hanser, C. (2007, March). *Tags are not metadata, but "just more content" to some people*. Paper presented at the International Conference on Weblogs and Social Media, Boulder, CO, USA.

Cheng, R., & Vassileva, J. (2005, January). *User motivation and persuasion strategy for peer-to-peer communities*. Paper presented at the the 38th annual Hawaii International Conference on System Sciences, Big Island, HI, USA.

Chi, E. H., & Mytkowicz, T. (2008, June). *Understanding the efficiency of social tagging systems using information theory*. Paper presented at the 19th ACM Conference on Hypertext and Hypermedia, Pittsburgh, PA, USA.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*, 627–668.

Deci, E. L., & Ryan, R. M. (1985). *Motivation and self-determination in human behavior*. Heidelberg, Germany: Springer.

Drenner, S., Shen, S., & Terveen, L. (2008, October). *Crafting the initial user experience to achieve community goals*. Paper presented at the the 2008 ACM Conference on Recommender Systems, Lausanne, Switzerland.

Golder, S. A., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science, 32*, 198–208.

Gould, J. D., & Lewis, C. (1985). Designing for usability: Key principles and what designers think. *Communications of the ACM, 28*, 300–311.

Heckner, M., Heilemann, M., & Wolff, C. (2009, May). *Personal information management vs. resource sharing: Towards a model of information behaviour in social tagging systems.* Paper presented at the International AAAI Conference on Weblogs and Social Media (ICWSM),San Jose, CA, USA.

Jäschke, R., Marinho, L., Hotho, A., Schmidt-Thieme, L., & Stumme, G. (2008). Tag recommendations in social bookmarking systems. *AI Communications, 21*, 231–247.

Lange, P. G. (2008). Publicly private and privately public: Social networking on YouTube. *Journal of Computer-Mediated Communication, 13*, 361–380.

Logan, R. A., Nuttall, R. J., Hazelwood, S. E., Parker, J. C., Johnson, J. C., Hewett, J. E. & Reid, J.C. (2000). Audience motivations to use an arthritis website. *Arthritis Care and Research, 13*, 320–329.

Marlow, C., Naaman, M., Boyd, D., & Davis, M. (2006, August). *HT06, tagging paper, taxonomy, Flickr, academic article, to read*. Paper presented at the Seventeenth Conference on Hypertext and Hypermedia, Odense, Denmark.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*, 370–396.

Mathes, A. (2004). *Folksonomies: Cooperative classification and communication through shared metadata*. Retrieved September, 10, 2010, from http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html

Matusiak, K. K. (2006). Towards user-centered indexing in digital image collections. *OCLC Systems & Services: International Digital Library Perspectives, 22*, 283–298.

Melenhorst, M., Grootveld, M., & van Setten, M. (2008). Tag-based information retrieval of educational videos. *EBU Technical Review, 314*, 1–18.

Melenhorst, M., & van Setten, M. (2007, month). *Usefulness of tags in providing access to large information systems*. Paper presented at the IEEE Professional Communication Conference, Seattle, WA, USA.

Morrison, P. J. (2008). Tagging and searching: Search retrieval effectiveness of folksonomies on the World Wide Web. *Information Processing and Management, 44*, 1562–1579.

Norris, C. E., & Colman, A. M. (1994). Effects of entertainment and enjoyment of television programs on perception and memory of advertisements. *Social Behavior and Personality, 22*, 365–376.

Nov, O., Naaman, M., & Ye, C. (2008, April). *What drives content tagging: The case of photos on Flickr*. Paper presented at the the 26th annual SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy.

Preece, J. (2000). *Online communities. Designing usability, supporting sociability*. New York: John Wiley & Sons, LTD.

Rashid, A. M., Ling, K., Tassone, R. D., Resnick, P., Kraut, R., & Riedl, J. (2006, April). *Motivating participation by displaying the value of contribution*. Paper presented at the CHI 2006, Montréal, Canada.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54–67.

Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F.M., & Riedl, J. (2006, November). *Tagging, communities, vocabulary, evolution*. Paper presented at the Conference on Computer Supported Cooperative Work, Banff, Canada.

Shang, R., Chen, Y., & Shen, L. (2005). Extrinsic versus intrinsic motivations for consumers to shop on-line. *Information & Management, 42*, 401–413.

Sood, S. C., Owsley, S. H., Hammond, K. J., & Birnbaum, L. (2007, March). *TagAssist: Automatic tag suggestion for blog posts*. Paper presented at the International conference on weblogs and social media, Boulder, CO, USA.

Stafford, T. F., Stafford, M. R., & Schkade, L. L. (2004). Determining uses and gratifications for the Internet. *Decision Sciences, 35*, 259–288.

Teo, T. S. H., Lim, V. K. G., & Lai, R. Y. C. (1999). Intrinsic and extrinsic motivation in internet usage. *Omega, 27*, 25–37.

Van Velsen, L., & Melenhorst, M. (2008, July). *User motives for tagging video content*. Paper presented at the Adaptation for the Social Web workshop, Hannover, Germany.

Van Velsen, L., & Melenhorst, M. (2009). Incorporating user motivations to design for video tagging. *Interacting with Computers, 21*, 221–232.

Von Ahn, L. (2006). Games with a purpose. *Computer, 39*, 92–94.

Von Ahn, L., & Dabbish, L. (2004, April). *Labeling images with a computer game*. Paper presented at the Conference on Human Factors in Computing Systems, Vienna, Austria.

Wartena, C. (2010). Automatic classification of social tags. In M Lalmas, J. Jose, & I. Frommholz (Eds.), *Lecture notes in computer science: Research and advanced technology for digital libraries* (pp. 176–183). Berlin, Germany: Springer-Verlag.

Wartena, C., & Brussee, R. (2008, September). *Topic detection by clustering keywords*. Paper presented at the the 19th International Conference on Database and Expert Systems Application, Turin, Italy.

Wu, H., Zubair, M., & Maly, K. (2006, August). *Harvesting social knowledge from folksonomies*. Paper presented at the the Seventeenth Conference on Hypertext and Hypermedia, Odense, Denmark.

Author's Note

We thank the students from the Utrecht University of Applied Science who participated in the software development and data collection for this study. The research was financed through the European Commission FP7 Project MyMedia (www.mymediaproject.org) under the grant agreement No. 215006.

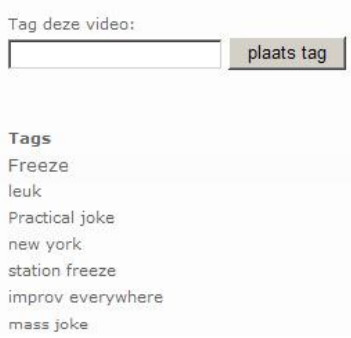
All correspondence should be addressed to
Mark Melenhorst
Novay
PO Box 589
7500 AN Enschede, the Netherlands
Mark.Melenhorst@novay.nl

Human Technology: An Interdisciplinary Journal on Humans in ICT Environments
ISSN 1795-6889
www.humantechnology.jyu.fi

APPENDIX A

Chatbot condition

Control condition



This is the control condition. In the text box, tags can be entered that appear in the tag list after pressing the *Plaats tag* button

Tag & Vote condition

Below *Uw gegevens* the user's statistics and position in the *higschore* is displayed. The green and red arrows represent positive and negative votes for tags. Below *Tag deze* video, new tags can be entered.

Bookmarking Condition

The user's collection of tags is displayed below *Mijn tags*. Folders can be found below *Mijn mappen*. In *Mijn favorieten* new video clips and tags can be entered.

A collage of screendumps of the tagging mechanisms.

APPENDIX B

Titles and URLs of YouTube Videos.

Title	URL
Frozen Grand Central	http://www.youtube.com/watch?v=jwMj3PJDxuo
Japanese way of folding t-shirts	http://www.youtube.com/watch?v=b5AWQ5aBjgE
Dove evolution	http://www.youtube.com/watch?v=iYhCn0jf46U
OFFICIAL - Terry Tate Office Linebacker "Superbowl Spot"	http://www.youtube.com/watch?v=Kg5cdZ-Fnpc
Learn Popular Magic Illusions : The Penetrating Pinky Illusion Magic Trick Explained	http://www.youtube.com/watch?v=RDGCR4W7Yn8
Fonejacker: Latest Episode: Bank Robber Vs Locksmith	http://www.youtube.com/watch?v=rr2d7YYUHEI
Bud Light Swear Jar	http://www.youtube.com/watch?v=JI3Y1auTFpU
Big Band Explained With Mince Pies	http://www.youtube.com/watch?v=WdCqtnS_cOA

APPENDIX C

After each movie clip the following questions were asked:

What do you think of this clip? Please indicate what you think of this clip on a scale from 1 to 5 on the following points.

Entertaining	1 – 2 – 3 – 4 – 5
Exciting	1 – 2 – 3 – 4 – 5
Humorous	1 – 2 – 3 – 4 – 5
Amusing	1 – 2 – 3 – 4 – 5
Nice	1 – 2 – 3 – 4 – 5
Funny	1 – 2 – 3 – 4 – 5

What mark between 1 and 10 would you assign to this clip?

Suppose you would see this clip on a video site. Please indicate to what extent you would like to provide this clip with tags.

- I would definitely not tag this clip
- I would probably not tag this clip
- I am not sure if I would tag this clip
- I would probably tag this clip
- I would definitely tag this clip

Apart from questions about the movie clip, after the second movie of each pair associated with a tagging condition, the following questions were asked about the tagging input mechanism:

Please provide a mark between 1 and 10 for this interface:

Please indicate for the statements below to which extent you agree with them:

This interface is easy to use	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
It is easy to learn how to use this interface	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
When I saw the interface, I could see what I could do immediately	Strongly disagree – Disagree – with it Neutral – Agree – Strongly agree
The interface is fun to use	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
If this website would exist in real life, I would definitely use it	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
The interface encourages to tag	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
Tagging has added value for me	Strongly disagree – Disagree – Neutral – Agree – Strongly agree