A Cognitive Perspective on Emergent Semantics in Collaborative Tagging: The Basic Level Effect

Tobias Ley[1,2], Paul Seitlinger[2]

[1] Know-Center[1], Inffeldgasse 21a, 8010 Graz, Austria,
tley@know-center.at
[2] Cognitive Science Section, University of Graz, Universitätsplatz 2, 8010 Graz, Austria
paulchristian.seitlinger@edu.uni-graz.at

Abstract. Researching the emergence of semantics in social systems needs to take into account how users process information in their cognitive system. We report results of an experimental study in which we examined the interaction between individual expertise and the *basic level advantage* in collaborative tagging. The basic level advantage describes availability in memory of certain preferred levels of taxonomic abstraction when categorizing objects and has been shown to vary with level of expertise. In the study, groups of students tagged internet resources for a 10-week period. We measured the availability of tags in memory with an association test and a relevance rating and found a basic level advantage for tags from more general as opposed to specific levels of the taxonomy. An interaction with expertise also emerged. Contrary to our expectations, groups that spent less time to develop a shared understanding shifted to more specific levels as compared to groups that spent more time on a topic. We attribute this to impaired collaboration in the groups. We discuss implications for personalized tag and resource recommendations.

Keywords: Tagging, Categorization, Personalized Recommendation

1 The Basic Level Effect in Collaborative Tagging

Emerging semantics in social systems is a topic that has sparked significant interest in the research community. In collaborative tagging systems [5], for example, it has been suggested that a community of users negotiates meaning in a collaborative sensemaking process [4] that would lead to a stabilization of the used vocabulary over time [5]. Some have suggested that this process could be an alternative to the usually top down driven engineering of ontologies [3] [8] [11]. An example for this is the Software SOBOLEO (Social Bookmarking and Lightweight Engineering of Ontologies [1])

[1] The Know-Center is funded within the Austrian COMET Program (Competence Centers for Excellent Technologies) under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG. Work on this paper has been partially funded within the MATURE project (www.mature-ip.eu) by the European Community.

that was used in our study. Here, users collaboratively tag bookmarks and then use the tags to build a shared vocabulary and a taxonomic structure.

Our conjecture is that besides an understanding of the social (e.g. [2]) and pragmatic processes (e.g. 6], it is equally important to understand the underlying cognitive processes in collaborative tagging for offering effective recommendations. For example, it has been shown that human categorization processes are highly variable and adaptive. Categorization does, for instance, vary on the level of specificity depending on a number of factors. Therefore, our intention with the study reported here is to look at the temporal dynamics in the collaborative tagging environment both in terms of the tagging activities and the associated cognitive processes over time. By doing so, we would like to gain a better understanding of the variability in human categorization as it can be observed in such an environment, and thereby enhance current personalized tag recommendation mechanisms provided both in the process of tagging and in the process of browsing tag clouds and resource collections. This should enhance the emergence of stable patterns in these environments.

The term *basic level advantage* [10] has been introduced to describe a preferred level of taxonomic abstraction when classifying objects of the real world (e.g. a preference for the term "dog" as opposed to "mammal" or "poodle"). In human communication, the basic level has an important role as it contains categories that are most easily retrieved from memory and have a high degree of information value in describing objects. Among many others, an advantage for the basic level has been shown when people verify the categories of pictures of objects [10], or in a free naming paradigm [12]. While the role of the basic level advantage in collaborative tagging is often acknowledged [1] [5] [6], surprisingly little empirical research exists to inform design decisions. In their study of *delicious*, Golder & Huberman [5] suggest that popular tags which are introduced very early for a certain bookmark correspond to categories of the basic level. The authors also find that the tag distribution for a certain bookmark quickly stabilizes over time suggesting an emerging consensus.

The authors also point to a potential problem with the basic level advantage that arises with differing levels of expertise. They hypothesize that there should be systematic variations across individuals of "what constitutes a basic level". In collaborative tagging, this basic level variation is a potential drawback. When resources are described on varying levels of specificity, it makes retrieval of information more difficult both for experts and for novices. While for the former, the information value of a basic level category is too low, for the latter the specific categories are not sufficiently well represented in memory, and, hence, their labels difficult to comprehend.

The hypothesized basic level variation is in line with cognitive research which has found a basic level shift in various categorization paradigms, such as generating attributes of category objects, free naming of category labels or verifying category membership [9] [10] [12]. Basic level shift for more experienced persons leads to better availability in memory of category members and their attributes on more specific levels of the taxonomy. Following sensemaking research, we expect that in a collaborative tagging environment a growing expertise in the domain can be observed over time. Therefore, we hypothesize that users will use more specific categories, will show better availability of these in memory and will ascribe more importance to more specific categories, when they collaboratively tag for a longer as compared to a shorter duration of time.

2 An Experimental Study

To test this hypothesis, we asked four groups of students to collaboratively collect bookmarks related to their course subject and describe them with tags. Two of the groups had to work on a topic for the whole duration of the semester (10 weeks), the other two groups switched their topic at half time. Our hypothesis was that the *long duration (ld)* groups would form a stronger representation in memory of the more specific tags and that they would rate their relevance higher than the *short duration (sd)* groups. Collaborative tagging among the students was realized through the social bookmarking system SOBOLEO. In SOBOLEO, the tags and the tag taxonomy that is collaboratively created are shared by all users of the system.

2.1 Participants and Procedure

The study took place in the context of a university course on cognitive models in technology enhanced learning at the University of Graz. Subjects (N=25, mean age M=23.3, SD=1.2) were psychology students participating for course credit. After an introduction to SOBOLEO, a computer literacy questionnaire and a word association test eliciting participants' knowledge about central concepts of the given topics were administered to the participants. Subjects were then assigned to four groups of 6 or 7 participants which were equivalent according to their scores on the word association test and computer literacy questionnaire. Each group was provided with their own SOBOLEO instantiation only accessible by personal usernames and passwords.

E-mails were then sent out to inform the participants of the topic they had to work on together with access details for their SOBOLEO environment. Two groups were asked to research the topic "the use of Wikis in enterprises", the other two groups "the use of Weblogs in universities". They were asked to prepare these topics as if they were collaboratively working on a report of presentation. Both topics were chosen because they were related to the course subject and because we expected the participants to have only little prior knowledge about them.

During the whole duration of the study (ten weeks) each student was expected to post two relevant bookmarks per week to the SOBOLEO environment and describe them with meaningful tags. The students were also required to collaboratively organize their tag collection with the help of the SOBOLEO taxonomy editor. To facilitate the emergence of consensus, the students were also encouraged to utilize the SOBOLEO chat and an external discussion forum.

After five weeks (at halftime), the SOBOLEO environments of two of the four groups were cleared. They had to start from scratch and to work on the other topic for another five weeks, making them the short duration (*sd*) groups. The other two groups continued with their prior topic, making them the long duration (*ld*) groups. Right before this topic switch, we also controlled for the fact that there still were no differences between the two conditions in the word association test. At the end of the semester, the association test and the relevance rating were administered to the 25 students in a group setting using a sample of tags they had created so far.

2.3 Tag Samples, Tag Specificity and Dependent Measures

By the end of the 10-week period, the four groups had created N=213 tags from which n=76 tags were drawn as a sample. To yield the independent variable *tag specificity,* tags were drawn from three different levels of the SOBOLEO taxonomies the students had created: *General* tags were drawn from the taxonomy levels 1, *medium* tags from level 2, and all tags below level 2 were allocated to the *specific* tags. From each of the four SOBOLEO environments, 19 tags were randomly drawn: three general (e.g. "weblogs", "e-learning by collaborating"), eight medium (e.g. "kinds of weblogs", "psychology of weblogs") and eight specific tags (e.g. "videoblogs", "microblogging"). Hence, the entire sample consisted of 76 tags: 12 general, 32 medium and 32 specific tags.

As a dependent measure, a relevance rating was collected at the end of the semester asking subjects to rate each tag sampled from their own SOBOLEO environment on a five-point Likert scale ranging from strongly relevant to strongly irrelevant for describing and organizing resources. By averaging the ratings of all group members a mean relevance rating for each tag was obtained. An association test was also conducted at the end of the semester. This test elicits implicit knowledge about concepts underlying verbal representations. Subjects were confronted with tags as stimulus words and asked to write down all associations coming to their mind. Response time was confined to 30 seconds. By counting the number of associations, the test informs about the strength of representation of concepts in memory. Stimulus words were the same tags used for the relevance rating. Again we averaged the number of associations of all group members to obtain a mean number of associations for each tag.

3 Results

Figure 1 displays the mean *number of associations* (left) and the mean *relevance rating* (right) as a function of tag specificity and duration obtained at the end of the study. These results indicate a basic level advantage, i.e., a strong representation of categories represented by general tags. Independent of duration, general tags at level 1 seem to evoke more associations (M=4.43, SD=0.21) than medium tags at level 2 (M=2.95, SD=0.13) and specific tags at level 3 (M=2.99, SD=0.13).

Secondly, a level - group interaction is emerging, but it is in the opposite direction than we had expected. Contrary to our expectations, *sd* groups achieved more associations and higher relevance ratings than *ld* groups for medium and specific tags. This was confirmed by a duration (*ld* and *sd*) × tag specificity (*medium* and *specific*) multivariate analyses (MANOVA) on the variables *number of associations* and *relevance rating*. The main effect for *duration* proved highly significant ($F_{2,58}$=9.82, p<.01) explaining 25% of variance in the dependent variables and indicating a strong effect ($\hat{f} > .40$). Neither the main effect *tag specificity* nor the interaction between *duration* and *tag specificity* were significant. To further determine the nature of the significant effect, two univariate analyses (ANOVAs) for each of the dependent variables were conducted. Both results match our descriptive pattern. Averaging over *medium* and *specific* tags, the *sd* groups achieve more associations (M=3.29, SD=0.61) than the *ld*

groups (M=2.65, SD=0.81; $F_{1,59}$=13.01, p<.01). The same applies to the relevance rating ($F_{1,59}$=9.12, p<.001): the judged relevance of *medium* and *specific* tags is higher in *sd* groups (M=2.56, SD=0.60) than in *ld* groups (M=2.22, SD=0.69).

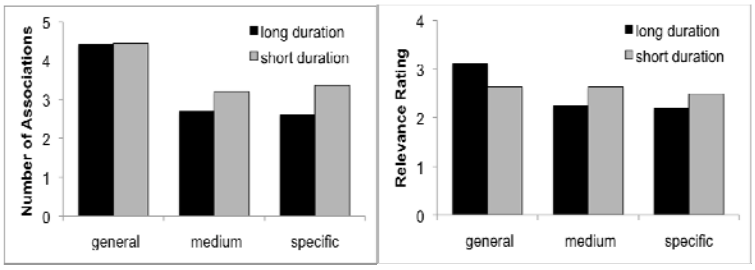

Fig. 1. Number of associations (left) and relevance rating (right) for general, medium and specific tags in long duration (10 weeks) and short duration (five weeks) groups

Results of a post-hoc questionnaire that had been administered to the students at the end of the semester give insight into these counterintuitive findings. First, all groups indicated they had been dissatisfied with the communication mechanisms (the SOBOLEO Chat and discussion forum). Albeit having worked on their topic for a longer time, groups of the *ld* condition gave significantly lower ratings when asked for the understanding of the topic (M=1.67 on a 5-point Likert scale, SD=1.23) than *sd* groups (M=2.69, SD=0.75; $F_{1,23}$=6.44, p<.05). Additionally, *ld* groups (M=1.92, SD=1.00) perceived a lower quality of their taxonomy than *sd* groups (M=2.92, SD=0.86; $F_{1,23}$=7.33, p<.05). Free text answers indicate that especially students in *ld* groups found it more difficult to collaboratively work on the shared taxonomy in SOBOLEO and they felt that the exercise had resulted in a chaotic collection of bookmarks and tags where it was rather difficult to keep an overview.

4 Discussion and Outlook

We conclude from the study that a strong basic level effect could be observed for an implicit memory measure (number of associations) as well as an explicit measure (relevance rating), where for the latter this only showed for one of the groups. However, our manipulation (duration of engagement with a topic) was obviously not effective in producing a stronger representation in memory. Quite to the contrary, the fact that environments of students in *sd* groups were cleared after half time actually helped them to build a more effective and shared external knowledge representation. The negative effect for *ld* groups was exacerbated by missing effective communication mechanisms in the SOBOLEO system. Similarly, we assume that it was students from *sd* groups that developed a more shared and stronger internal representation. If this was the case, then there is clear evidence for a shift in the basic level. This already showed after a comparatively little time (5 weeks), and produced a strong and also practically significant effect (an increase of 0.64 associations on average).

Results of this study have practical significance for tag and resource recommendation in collaborative environments (e.g. [2]) as they suggest that effective tag recom-

mendations need to take tag specificity into account. Experts in a domain would benefit from more specific tag recommendations or from recommendations of resources with more specific tag assignments. The study also suggests that temporal dynamics need to be taken into account where shifts in basic level already take place after a few weeks of collaboration. Finally, in case tag specificity could be captured, this would also have implications for user modelling as the level of expertise pertaining to a certain topic could be derived for any user from his or her tag assignments.

A limitation of our results relates to the manual creation of the taxonomy by students which extends the (normally flat) folksonomy by a taxonomic relation. For our future work, we plan to draw on statistical approaches, such as [3] who found different tag similarity measures (tag co-occurrence vs. distributional measures) to correspond to different taxonomic relationships between tags. Moreover, these results seem to be moderated by particular behavioural tendencies of users using the tagging system [6].

References

1. Braun, S., Schmidt, A., Walter, A., Nagypal, G., Zacharias, V.: Ontology Maturing: a Collaborative Web 2.0 Approach to Ontology Engineering. In: 16th int'l WWW conference, pp. 1-10 (2007).
2. Carmagnola, F., Vernero, F., Grillo, P.: SoNARS: A Social Networks-Based Algorithm for Social Recommender Systems. In: 17th int'l UMAP conference, pp. 223-234 (2009).
3. Cattuto, C., Benz, D., Hotho, A., Stumme, G.: Semantic Grounding of Tag Relatedness in Social Bookmarking Systems. In: 7th ISWC Conference, pp. 615-631 (2008).
4. Fu, Wai-Tat.: The microstructures of social tagging: a rational model. In: Proceedings of the ACM 2008 conference on Computer supported cooperative work, pp.229-238 (2008).
5. Golder, S., Huberman, B.A. The Structure of Collaborative Tagging Systems. J. Information Sciences 32, 198-208 (2006).
6. Körner, C., Benz, D., Hotho, A., Strohmaier, M., Stumme, G.: Stop Thinking, Start Tagging: Tag Semantics Emerge From Collaborative Verbosity. In: 19th Int'l WWW Conference, ACM, New York, 2010.
7. Marlow, C., Naaman, M., Boyd, D., Davis, M: HT06, tagging paper, taxonomy, Flickr, academic article, to read. In Proc. HYPERTEXT '06, ACM Press, 31-40 (2006).
8. Mika, P.: Ontologies Are Us: A Unified Model of Social Networks and Semantics. J. Web Semantics: Science, Services and Agents on the World Wide Web 5, 5-15 (2007).
9. Rogers, T.T., Patterson, K.: Object categorization: reversals and explanations of the basic-level advantage J. Experimental Psychology: General 136, 451-469 (2007).
10. Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., Boyes-Braem, P.: Basic objects in natural categories. J. Cognitive Psychology 8, 382-439 (1976).
11. Specia, L., Motta, E.: Integrating Folksonomies with the Semantic Web. In: 4th ESWC2007 European Semantic Web Conference, pp. 624--639. Springer, Heidelberg (2007).
12. Tanaka, J.W., Taylor, M.: Object Categories and Expertise: Is the Basic Level in the Eye of the Beholder? J. Cognitive Psychology 23, 457--482 (1991).