Designing and Evaluating Kalas: A Social Navigation System for Food Recipes

MARTIN SVENSSON, KRISTINA HÖÖK, and RICKARD CÖSTER
Swedish Institute of Computer Science

The idea of social navigation is to aid users to navigate information spaces through making the collective, aggregated, or individual actions of others visible and useful as a basis for making decisions on where to go next and what to choose. These social markers should also help in turning the navigation experience into a social and pleasurable one rather than the tedious, boring, frustrating, and sometimes even scary experience of a lonely traveler. To evaluate whether it is possible to design for social navigation, we built the food recipe system *Kalas*. It includes several different forms of aggregated trails of user actions and means of communication between users: recommender system functionality (recommendations computed from others' choices), real-time broadcasting of concurrent user activity in the interface, possibilities to comment and vote on recipes, the number of downloads per recipe, and chatting facilities. Recipe author was also included in the recipe description.

Kalas was tried with 302 users during six months, and 73 of the users answered a final questionnaire. The overall impression was that users liked and acted on aggregated trails and navigated differently because of them. 18% of the selected recipes came from the list of recommended recipes. About half of the 73 users understood that recommendations were computed from their own and others actions, while the rest had not reflected upon it or had erroneous beliefs. Interestingly, both groups selected a large proportion of their recipes from the recommendations.

Unfortunately, there were not enough users to populate the space at every occasion, and thus both chatting and following other users moving in the space was for the most part not possible, but when possible, users move to the space where most other users could be found. Of the other social textures, users themselves claimed to be most influenced by other users' comments attached to the recipes and less by recipe author or number of downloads. Users are more positive to the possibility of expressing themselves in terms of comments and voting than seeing the comments and votes of others.

It was noted that users did not pick more recommended recipes towards the end of the study period when the accuracy of recommendations should have been higher. More or less from the start, they picked recommended recipes and went on doing so throughout the whole period.

Categories and Subject Descriptors: H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Information filtering*; H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces—*Collaborative computing*; H.5.4 [**Information Interfaces and Presentation**]: Hypertext/Hypermedial—*Navigation*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evaluation / methodology*

This research was funded by The Swedish IT Institute (SITI).
Authors' address: M. Svensson, Swedish Institute of Computer Science, 164 29 Kista, Sweden; email: martins@sics.se.

General Terms: Design, Experimentation, Human Factors

Additional Key Words and Phrases: Social navigation, collaborative filtering, experiments, evaluation, design

1. INTRODUCTION

Net-based services have some unique possibilities of aiding users that until now have remained largely unexplored namely, the presence of a large number of other users. These other users may already have explored the application is functionality, they may have domain knowledge that could aid others, and, perhaps even most important, they may have opinions and subjective evaluations of what the application offers that could guide or influence others. But how can this wealth of social texture be collected and accumulated to actually aid users navigation through space without disturbing them?

This is the core problem of the field of social navigation [Dieberger et al. 2000; Munro et al. 1999; Höök et al. 2002]. The benefits of social navigation are not only to accumulate users' trails and by these guide others in navigating large information spaces. The benefits also include giving users a sense of social presence, of not being alone in the space. It may provide users with a subjective stance—a texture of others' judgments that will aid in their choice of whatever the application offers be it music, movies, books, or as in the present case, food recipes.

While social navigation ideas have already spread from the research community to the design of commercial systems such as amazon.com, only a few end-user studies (e.g., Wexelblat [2002], Svensson et al. [2001]) exist that establishes how and when social navigation can be added to a system. The study of Kalas aims to further our understanding of two different but related issues:

—*Effects on navigation*. Is it the case that social navigation will aid users to navigate more efficiently? How can this be evaluated?

—*Perceived subjective quality*. Does social navigation provide a social texture that adds to the perceived quality of the navigational process and choices made (or even the application itself)?

To evaluate these issues, a stable fully-functioning system, Kalas, which could be tested in a real situation competing with other commercial solutions, was built. Kalas was then exposed to 302 active users for six months. The logs of usage during the six months, the final questionnaire with 73 of the users, and in-depth interviews with four users, gave an overall impression that users liked the social parts of Kalas even if not everyone acted on all of them. 18% of selected recipes came from the list of recommended recipes. About half of the 73 users understood that recommendations were computed from their own and others actions, while the rest had not reflected upon it or had erroneous beliefs. Interestingly, both groups selected a large proportion of their recipes from the recommendations. The system managed to convey the meaning of a "thumbs-up" symbol so that users did, in fact, understand that the recommendations came from other users' choices.

One reason for not navigating the recipe space by what was said in chat or the real-time presence of others was that there were not enough users to populate the space at every occasion. Still, on those occasions when it was possible, it seemed as if users were influenced by others real-time presence and movements in, the space (when five or more users were logged in, most users went to the more populated parts of the space). Twelve also tried to chat.

Of the other social textures, users themselves claimed to be most influenced by other users' comments attached to recipes and less by recipe author or number of downloads. Interestingly, users are more positive about the possibility of expressing themselves in terms of comments and voting than seeing the comments and votes by others.

Before we go into the study and how we arrived at these conclusions, we provide a short introduction to the idea of social navigation and a description of Kalas.

2. SOCIAL NAVIGATION

Dourish and Chalmers [1994] introduced the concept of social navigation. They saw it as "navigation towards a cluster of people or navigation because other people have looked at something." In parallel with their work, Hill et al. [1992] introduced the idea of edit wear and read wear. By tagging scrollbars with read and edit patterns, they created the first history-enriched environments. Around this time, collaborative filtering or recommender systems (e.g., Shardanand and Maes [1995] and Konstan et al. [1997]) started to become popular. By collecting the opinions of a large number of people, an individual can specify items that they like or dislike, and then the system recommends other items based on the data collected from other people.

Later, Dieberger [1997] widened the scope in seeing direct recommendations of, for example, Web sites and bookmark collections as a form of social navigation. He was inspired by the remarks made by Erickson [1996] that the Web could be characterized as a social hypertext where nodes represent people. The links, as well as the page itself, provide a view of a person's network of friends, colleagues, and interests.

Erickson et al. [1999] later went on to work ideas for *socially translucent systems* and *social proxies*. Socially translucent systems reciprocally show the activities of users so that they can be held accountable for their actions. It is a generic perspective on how to deal with the problem of privacy in different social systems, including social navigation systems. If you can see my actions, and you know that I know that you see me, I can be held accountable for what I do. Erickson et al. [1999] applied these ideas to their slant on social navigation systems, focusing in particular on the real-time presence of users: social proxies. In the system Babble, Erickson et al. implemented the social proxy concept. Babble is a computer-mediated communication system in which users are represented as marbles within a circle. The marbles, colors and positions shift depending on whom talks to whom and who is active [Erickson 2004].

Fundamental to Social Navigation is the observation that much of the everyday information seeking is carried out through watching and talking to other people. When navigating cities, people tend to ask other people for advice rather than study maps [Streeter et al. 1985]; when trying to find information about pharmaceuticals medical doctors ask other doctors for advice [Timpka and Hallberg 1996]. Munro [1999] observed how people followed crowds or simply sat around at a venue when deciding which shows and street events to attend at the Edinburgh Arts Festival.

However, observing that social navigation happens in the world does not necessarily mean that it is a good idea to design systems from this perspective. Social navigation cannot be unproblematically translated into ready-made algorithms and tools to be added on top of an existing space. What can be done is to make information spaces afford social interactions and accumulate social trails. Social navigation is a dynamic interaction between the users, the items (whether food recipes, books, or something else), and the activities in the space. All three are subject to change over time.

Moreover, users have to conceptualize representations such as sets of stars as in Movielens [Herlocker et al. 2000] (see Figure 1), pick-and-pop[1] numbers at download.com, or the footprints next to links in the socially enhanced SWIKI system (see Figure 1) [Dieberger and Lönnqvist 2000] as being representations of what other users have done before them. Oftentimes, this turns out to be difficult. Studies of alternative explanation models tested in the Movielens system have shown that users are not necessarily helped by rich explanations telling the full story of where recommendations come from [Herlocker et al. 2000].

A range of systems has been implemented that exhibit some of the social navigation properties. The most well-known commercial example is the Amazon product recommender: "others who bought this book also bought...". Research laboratory work includes the Footprints system [Wexelblat and Maes 1999] that visualizes history-enriched information. Similar ideas are explored in IBM's WebPlaces [Maglio and Barrett 1999]. It observes peoples' paths through the Web and looks for recurring paths.

3. KALAS—SOCIAL NAVIGATION OF FOOD RECIPES

Food recipes are particularly suitable as a domain for social navigation for several reasons. A typical recipe collection found on the Internet contains thousands of recipes that can be difficult to navigate (in Kalas there are over 3000 recipes to choose from). In this respect, the domain offers an interesting navigation problem.

Our choices of food are, to a large extent, driven by our individual taste which could be difficult to express solely in terms of recipe ingredients or some other content-based information. A recommender system, on the other hand, can be used to model users' taste by clustering users into groups that share similar taste and then give recommendations based on the clustering.

Cooking and choosing what to cook is often a social activity. People like to talk about what to cook, ask their friends about new interesting dishes, or turn to their favorite recipe author for recommendations. Cooking TV shows are extremely popular and people grow to trust some chefs more than others.

[1]Pick is the experts' choices, while pop reprsents popular downloads.

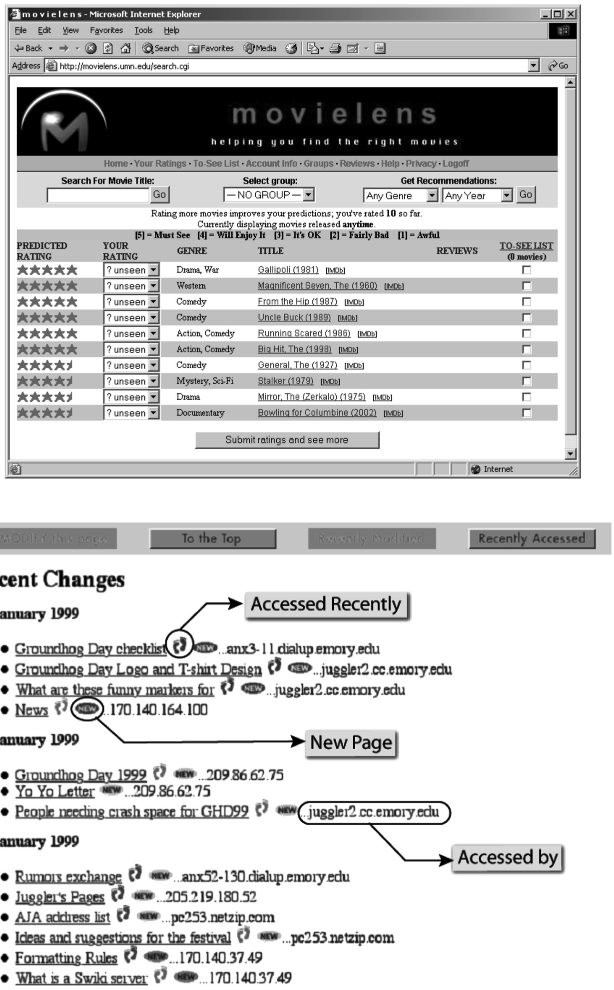

Fig. 1. The interfaces of Movielens (left), using stars for voting and presentation of recommendations, and the modified SWIKI system (right), using small footprints in different colors to represent how hot the trail of others for a particular link is.

People often want to develop their cooking skills and explore new kinds of food. Again, a scenario in which social navigation can be of help since it allows users to find information (or recipes) they did not explicitly ask for [Shardanand and Maes 1995].

Our goal was to turn the Kalas interface into a surface upon which various different social trails could be accumulated and displayed. We wanted the organization and presentation of the recipes to reflect user activities. By dividing the recipe space into smaller recipe collections with specialized themes (Italian, Mediterranean, Chef's choice, etc.), users can move between them and get different recommendations in different collections. Thus, finding recipes is a multistep process that typically entails first moving between recipe collections, then selecting recipes from a scrollable list, choosing a particular recipe,

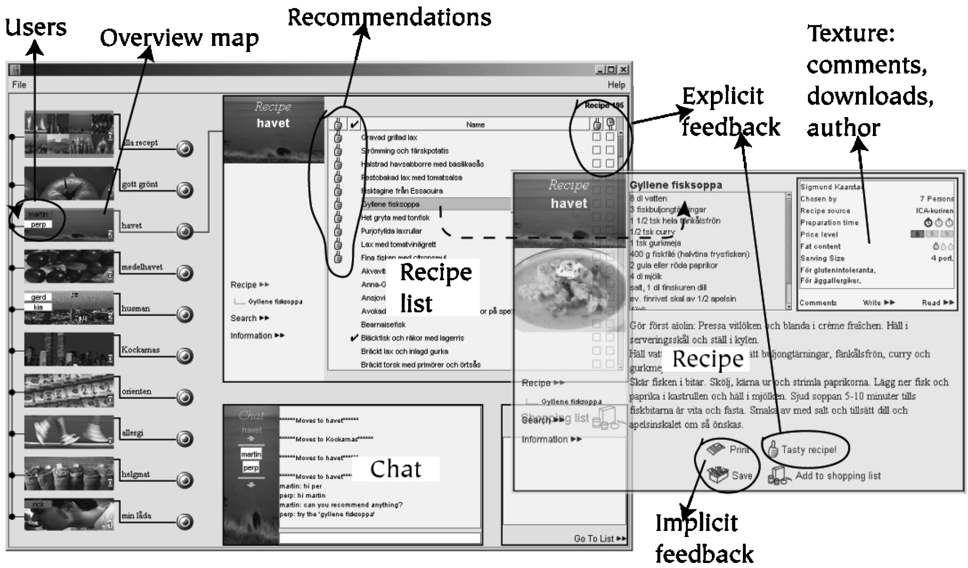

Fig. 2. The Kalas interface: once the user clicks on a recipe in the list, the area with recipes is replaced by the individual recipe, including user comments and other social trails.

studying its details, and then making a choice and/or going back over any of the previous steps.

A research goal was to design for different kinds of social navigation in order to study their respective effects on user behavior. Thus, we added social navigation markers to each of the navigational steps. Moving between recipe collections may be influenced by how many users are logged into each collection—the real-time presence of others. The recommender functionality affects which recipes turn up among the first ten recipes in each recipe collection. Each recipe may have comments attached from other users, and the number of times it has been downloaded will be shown. Another way of finding recipes is via chatting with other users in the chat rooms.

3.1 Kalas Design

To minimize design errors affecting our evaluation, we did a qualitative design study of a prototype version of Kalas. It was a controlled lab study conducted on a small set of subjects (12 in total) who were computer science students [Svensson et al. 2001]. Kalas was designed based on the experiences from the study, and the final design is shown in Figure 2.

The Kalas interface consists of a navigation overview with the recipe collections, a recipe list, a chatting area, a shopping list, and finally individual recipes. In each of these, social trails are shown. Within each recipe collection currently logged on, users are displayed in an overview map, recommendations based on others' choices of food recipes are shown in the recipe list, social texture such as comments, number of downloads, and recipe author is attached to the recipes, and, finally, real-time communication between users is available in the chat area.

Online awareness. Users navigate by moving between different recipe collections. A collection contains a fixed amount of recipes with a specific theme (e.g., vegetarian or spicy food). Users that are currently logged on can see each other in the navigation overview area and follow each other's movements between recipe collections. Users are displayed as labels with their names. The overview map is the main vehicle to aggregate online awareness.

Implicit and explicit feedback. A recommender needs some way of knowing which items a user prefers. Kalas allows users to provide feedback on recipes in two ways.

—*Implicit feedback*. Printing, saving, or adding a recipe to the shopping list will render a positive vote for that recipe.
—*Explicit feedback*. Clicking on the button "good recipe" underneath the recipe or checking the thumbs-up/thumbs-down option in the recipe list will render an explicit positive/negative vote for that recipe.

A user's explicit and implicit feedback for a recipe is grouped together in a single recipe rating. We used $-1/+1$ for explicit negative/positive feedback and 0.5 for implicit feedback. A negative/positive feedback overwrites any previous rating, whereas the implicit rating is only used if there was no previous feedback for that recipe. For the default value, Breese et al. [1998] recommends a neutral or slightly negative rating. We decided to take a slightly positive value (0.1) instead, due to the high quality of the recipes.

Recommendations. Whenever a user moves to a new recipe collection, the list of recipes in that collection is displayed. The first ten recipes are tagged with a thumbs-up symbol that signals that the recipes are recommended by the system. Consequently, users do not explicitly ask for recommendations but are always given ten recommendations whenever they move to a new recipe collection.

Users can at any time change the recipe list by sorting it in either alphabetical order, by chosen recipes, by recommendation, or by date. It is also possible to search for recipes in each collection through a keyword-based search system.

The recommendations are calculated by a memory-based collaborative filtering algorithm taken from Breese et al. [1998]. The recommended value of an item (recipe) is the active user's (the user receiving the recommendation) mean rating value, plus a weighted sum of other users' ratings for that recipe. The weight between the active user and another user is taken to be the Pearson correlation between the two users' rating vectors. Since the Pearson correlation will only match recipes that both users have rated, we used the extension of default ratings. A default rating is used when one of the users have rated a recipe, while the other has not. This extension enables calculating the correlation in the union of the rating vectors and helps alleviate the problem of sparsity. Another extension is the inverse user frequency where high frequency items are given a lower weight in the belief that items that many users have rated are not as useful for measuring user similarity as low frequency items. This extension was not used since we had little reason to believe that the recipe domain would exhibit the same frequency trends as, for example, the movie domain, where a small number of movies is seen by a large number of users.

Finally, we performed a very small bootstrapping exercise that generated ten artificial users who rated a few recipes according to a keyword search among recipes.

Social texture. The social textures of recipes consists of comments made by users, the number of times it has been chosen, the source from where the recipe originated (a company, other users, or some other source), and the recipe author. In Kalas, a food company produced the main bulk of the recipes, but a few recipes were provided by end users.

Communication. Users communicate synchronously in chatting areas connected to each recipe collection. Users can also leave comments on individual recipes that allow them to communicate asynchronously.

Invisibility. Kalas was designed to allow users to be invisible to others at the price of loosing some of the social features. Invisible users can see where other (visible) users are but lose the ability to participate in the public chat room. Kalas offers three visibility modes: visible, invisible, or "marked as friend". These different states are shown through the labels in the recipe collection area: friends are marked with a different color label.

4. EVALUATING SOCIAL NAVIGATION

Social navigation requires users. A study with few users can be useful in informing design—which is why we performed our initial study as part of the design and redesign of Kalas—but to study the impact of social navigation, the system has to be used by many users over a longer time period. Only after some time of usage will users move on from barely noticing subtle social cues to actually using them to empower their usage.

Before exposing Kalas to many users during a longer time period, it was necessary to decide how successful social navigation would be so that the important actions were logged and the right questions asked. Evaluating social navigation is difficult. It is usually not possible to determine whether a navigational step or choice made by a user is done because of the social trails left in the interface or because the user liked the item anyway. In a food recipe system like Kalas, there are many reasons for choosing a particular food recipe; is can be chosen because of its ingredients, the photograph of the dish, or because the social trails lead to it. Log statistics can only show whether the recommended recipes were chosen more often than other recipes and whether users moved to the most populated recipe collections. Questionnaire replies can tell us something more of whether users saw and appreciated the social functions. Interviews can provide us with a more in-depth understanding of the perceived quality. This is why we decided to collect several different kinds of data in the study of Kalas.

As discussed in the introduction, our study of Kalas focused on two different issues:

—*Effects on navigation*. Is it the case that social navigation will aid users in navigating more efficiently? How can this be evaluated?

—*Perceived subjective quality*. Does social navigation provide social texture that adds to the perceived quality of the navigational process and choices made (or even the application itself)? This requires that the social navigation design is noticed and understood by users and that they act based on it.

4.1 Effects on Navigation

In choosing food recipes, users make a number of conscious and unconscious reflections based on the recipe's ingredients, any photographs of the dish, how complex it will be to cook, its popularity, and so on, before deciding which recipe to choose. If the system works well, it is probably a combination of all these properties that makes a user choose a particular recipe. The social markers—real-time presence, recommendations, comments on recipes, and so forth—will only be one part of the decision-maxing process.

The effects of the real-time presence of others can be partly determined through the log statistics which shows where recipe-collection users move when several other users are logged in simultaneously. If they move more often to recipe collections with many users than to those with few or none, we might hypothesize that the presence of others affected their choice. Still, it might be that they would have moved there anyway which is why this needs to be complemented with questions about whether they felt influenced by others' movements.

How can we evaluate whether the recommender functionality improved the users' task of finding good recipes? One possibility is to use predictive accuracy metrics that measure how well the recommender algorithm can predict, from a sample set of votes, the future votes of a user. At first sight, it seems that such an evaluation could be appropriate. However, the predictive accuracy metrics tell us more about the underlying collaborative filtering algorithm than some qualitative aspects of the whole system. The correlation algorithm that we use has undergone empirical tests by others (e.g., Breese et al. [1998] and Calderón-Banavides et al. [2004]) and yet another evaluation of the accuracy of the algorithm is not of interest here. Instead, we are interested in the utility of the recommender.

In addition, testing the algorithms' predictive power implicitly assumes that an objective value can be assigned to the items (be it recipes, books, or movies) in the application. They are either right or wrong for users. Instead, it is sometimes the case that a lot of recipes fit more or less with users' taste and any of them can be chosen.

Recommender systems are decision-support systems; they should support the user in making a decision as to which (of many) items to select, purchase, download, and so forth. We believe that measuring whether the recommender functionality in Kalas did assist users in selecting recipes is more appropriate than measuring the error between a predicted rating and the actual rating. Cosley et al. [2002] describe a framework for evaluating recommender systems from the perspective of user acceptance rather than predictive accuracy. The framework utilizes statistics such as the number of times a recommended item is viewed or downloaded that is more appropriate for evaluating whether users act on the recommendations or not.

Whether the social trails on a particular recipe affect a user's decision to pick it or not cannot solely be determined from the log statistics. In addition, we have to rely on a users' questionnaire replies and the interviews.

4.2 Perceived Subjective Quality

Sometimes it is not enough that the information obtained from some system is relevant, it must also possess qualities that can only be determined from how other users react to it (the social texture). One recipe might look quite similar to another, but if someone tells you that one is better than the other, the quality of the recipes can be determined from trusting or distrusting the judgment of that person. In both cases, the subjective evaluation helps to determine the quality. It might even be that two recipes are equally tasty, have beautiful pictures, are roughly equally hard to cook, but if one is more popular, its quality can be perceived to be higher, and when chosen, users may feel more content and confident about their choice.

In studying how users react to recipes in Kalas, subjective quality can only be determined by what users themselves claim to be influenced by, and even then, they might not even be aware of why they act as they do. But a prerequisite for saying that the social trails in Kalas added to the perceived quality is if users at all *notice* and roughly understand the meaning of them. Second, once they have noticed and, in some sense, have understood their meaning, they should feel subjectively influenced by them. We attempted to capture these subjective aspects through some of the questions in the questionnaire as well as in the four interviews.

5. EVALUATING KALAS: METHOD

The evaluation took place over a period of six months, from June until November 2001, under as natural conditions as possible. Many different kinds of data were collected in the study including log data of activities in Kalas, questionnaires before and after completing usage of Kalas, and four in-depth interviews.

5.1 Subjects

The Kalas study was conducted in close collaboration with the online cooking portal hemma.net which provided us with over 3000 recipes. hemma.net is a Swedish Web site that has well over 10,000 active users. The portal mainly consists of information related to the home, for example, food recipes and refurbishment. The recruitment of subjects was carried out in two different ways. First, two emails were sent out to all subscribers at hemma.net. Second, a Web link to our description of the study and the system Kalas was added to the hemma.net portal. Thus, our subjects were from our target user group, that is, users interested in downloading recipes from the Web.

In all, we had 598 subjects initially signed up for the study. Most subjects, 309, were between 30 and 50 years old. Since subjects were recruited online, we got people from all over Sweden, mostly women living in smaller cities. 302 of these 598 users subsequently went on to actually install and use Kalas. This involved downloading a stand-alone and installing a Java-application. Lastly,

Table I. Subjects in the Kalas Study

	# Subjects	# Male	# Female	Mean Age	Use Internet Everyday	Skilled in Cooking	Computer Education
Signed up	598	68	530	50–59	445	213	225
Filled in final questionnaire	73	11	58	50–59	59	27	27

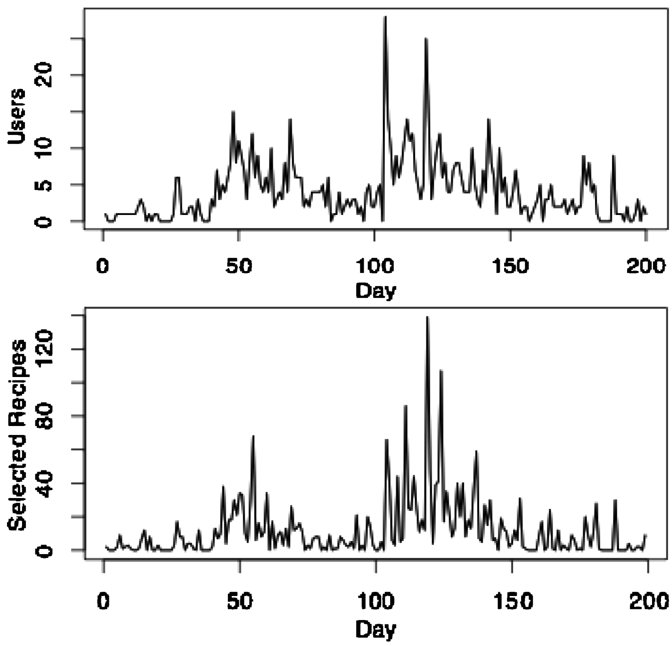

Fig. 3. Usage during six months: number of users logging on and number of selected recipes.

73 subjects answered the poststudy questionnaire. The groups are summarized in Table I.

5.2 Procedure

Subjects did not have any particular task to carry out and could use Kalas whenever they wanted and as much as they wanted. To use the system, subjects first had to fill in a prequestionnaire with background information (age, Internet experience, cooking experience, educational level). The recruitment period started in June when most Swedish people take their holiday, thus most users started using Kalas in August around day 50, see Figure 3. After approximately five months of use, subjects were prompted to fill in a postquestionnaire with a mixture of Likert-scale questions and statements to agree/disagree with. During the period of use all user actions were logged.

Four users were randomly selected for in-depth interviews to provide a qualitative evaluation of Kalas. While this was a small number of interviewees, the purpose of the interviews was to contextualize and better understand the findings in the postquestionnaire and conclusions drawn from log data, not to

Table II. Scores From 1 (Easy/Not Disturbing) to 7 (Difficult/Very Disturbing)

Postquestionnaire Question	1	2	3	4	5	6	7	Median (Mean)
How easy/difficult was Kalas to learn?	2	2	2	20	18	12	18	5 (5.1)
How easy/difficult was Kalas to use?	4	1	7	19	15	14	13	5 (4.8)

provide additional data. The in-depth interview, carried out over telephone, allowed users to express their subjective feelings towards the system.

6. RESULTS

We start by describing the average usage of Kalas and how that usage changed over the six months, and then describe the usage of four users interviewed to give an overview of how the system was used. This will provide some insights when we move on to evaluate effects on navigation and perceived subjective quality.

6.1 Learning and Using Kalas

Overall people found Kalas easy to learn (only six users found it more difficult than average) and use, see Table II.

Users understood the icons representing social presence and the voting by others fairly well. Of the 73 subjects,

—49 recognized the symbol for an online user;
—67 were aware of the chatting facility;
—70 knew that it was possible to comment on recipes;
—63 understood how to vote on recipes.

6.2 Usage

To better understand how active users were, how many times they logged on, how many recipes were viewed and chosen, three users have been extracted from the data.

The average active user. The average user is calculated as the mean of all actions from the 302 users. The average user used the system on 2.9 different days, visited 2.13 different collections per session, printed 1 recipe, saved 2 recipes, added 0.8 recipes to the shopping list, and gave explicit positive feedback on 0.97 recipes. Thus, in total, she voted (implicitly and explicitly) on 4.77 recipes. Furthermore, she turned invisible 0.17 times and edited her profile about 0.6 times (setting age, email, name, and description), and lastly, she visited around 15 recipe collections.

The average final group user. For the users who answered the final questionnaire, the average user looks slightly different. This user used the system for 5.8 days, visited 2.45 collections per session, visited 29 collections, and voted on a little more than 13 recipes (printed 2.8, gave explicit positive feedback on 3.15, saved 4.5, and added 2.2 to her shopping list). She edited her profile 1 time and was invisible 0.2 times.

Table III. Number of Actions on Recipes (Implicit votes for a recipe is when the recipe was saved, printed, or added to the shopping list in Kalas, while explicit votes are thumbs up/down actions)

Action	#
Recipes viewed	2315
Unique recipes viewed	789
Recipes chosen	1129
Unique chosen recipes	476
Recommended and chosen	199
Found through search	305
Visible but not recommended	90
Found through scrolling	535
Excluding search (1129-305)	824
Recipes that got explicit positive feedback	110
Recommended	31
Recipes that got explicit negative feedback	11
Recipes that got comments	11

The extreme user. The most active user is female; she used the system on 31 different days and logged in for 33 sessions in total but only visited 2.62 collections per session. She marked 17 other users as her friends and posted a total of 31 messages. This extreme user edited her profile over 60 times and changed collection 149 times during the period of use. This user voted on 20 recipes out of which 3 votes were explicit positive feedback.

6.3 Usage Over Time

Figure 3 shows how many people were logged in throughout the entire evaluation and the corresponding activity measured as number of viewed recipes. There are a couple of things to note. Apart from a few radical peaks, the number of users logged in is fairly stable. A second observation, crucial to some of our results, is that a recommender system will work better and better as it is given more votes. From a statistical perspective, the recommender system was not really able to predict users' preferences on an individual basis given that each user had, by the end of the study, on average only voted for 4.77 recipes. The importance and implications of this result will be discussed and analyzed in detail.

A community needs activity and when the number of users within a community is under a certain threshold, activity (or content) has to be provided from outside. The peaks in Figure 3 arise from outside interference. The peaks on day 48, 70, 105, 120, and 140 all coincide with emails sent to the users. Each activity peak is followed by a period of more activity. Thus, a moderator can actively make a system be more or less used [Girgensohn and Lee 2002].

Table III shows statistics on the number of recipes that have been viewed, selected, and commented.

6.4 Interviews

While the data in Table III provides for some understanding of how users made use of Kalas, the interviews provide another, more qualitative understanding

Table IV. Usage Data for In-Depth Interviewees

	Printed	Shopping List	Saved	Explicit Feedback	Personal Info	Comments, Chatted
Anna	16		2	3	X	X
Mia	1	3	4	5	X	
John	2		1		X	
Maria	3		4	3	X	

of what users really did and why they turned to the system. It also provides us with some insights into how much of the social textures users picked up, how they acted upon it, and its relative importance versus the recipe content itself. In the interviews, subjects' perspectives on issues like privacy, whether they thought that they could influence the system by their own actions or not, and views on chatting were discussed. Before the interview, a screendump showing a more densely populated instance of Kalas was sent to the person. Since there were rarely enough users logged in at the same time to provide for a real-time presence and influence, we wanted to know whether they would have felt influenced by the presence of others in the different recipe collections had their been enough users logged in at the same time.

Four subjects, named John, Maria, Anna, and Mia, see Table IV, were randomly chosen among the users logging in during the last month of the study to be interviewed over the phone.

Anna: 61 years old. Anna works with food and used Kalas because she enjoyed it and made use of it in her work. Anna used her real name as her log-in name but did not entered her email address. Among the profile data, she found the age information valuable to provide since it makes it easier for other users to determine whether they share a common ground if, for example, they want to chat.

Anna found the recipe list good and, as she explained, used it to search for recipes in her work. She mainly chose recipes based on ingredients. Anna wanted to read the comments by others on recipes but felt she was the only one who wrote comments. She liked the division into recipe collections but felt it lacked some additional categories filtering into, for example, desserts and main courses.

Anna had a fairly good idea of how the recommender system worked—her assumption was that it recommended recipes that other users had voted for. As a consequence, she has decided to provide explicit feedback (thumbs up/down) and she also saved recipes in "My Box". She claimed that she deliberately tried to influence the system to recommend the recipes that she had found interesting to others.

Being visible was not a problem to her. In fact, the only person she ever met in the system, she immediately tried to make contact with in the chat. She was, in general, very interested in making contact with users who shared her interests.

Mia: 25 years old. Mia used Kalas to search for new, interesting recipes. She used her real name as log-in (both first name and last name). She did not think about the consequences of giving out her real name, although she mentioned that experience from using other systems had made her sometimes regret it afterwards. In most cases, when an email address is required, she provides her hotmail address which is what she also did in Kalas.

Like Anna, she found the recipe list useful and the recommended recipes interesting. She found the division into recipe collections to be good, making it easy to find recipes.

Mia did not reflect on how Kalas chooses recipes to recommend. Still, being an active person, she had several times gone back to the system and voted on recipes she had cooked based on downloads in previous sessions. She also used explicit negative feedback, by marking recipes she believed would taste awful with the thumbs-down. When picking a recipe she usually studied the name and sometimes the recipe picture.

Mia used most functionality that Kalas offered; she saved recipes, made shopping lists, and read comments about recipes. Mia felt comfortable with being visible to others. On the other hand, no one had been in the system when she was logged on. If someone had been logged in simultaneously in Kalas, she would have moved to the appropriate collection and tried to start a discussion in the chat.

Mia said that the best part of Kalas was the simple interface and how easy it was to find the recipes she needed. She also found the interface appealing and the system fun to use. On the negative side, she felt that Kalas was somewhat difficult to use in the beginning. She got lost and had some difficulties in finding her way back.

John: 42 years old. John works professionally with food—as a chef—and used Kalas to search for recipes. When he used the system, he never saw any other users. He would have preferred a more traditional division of recipes in the system, with collections referring to meat, fish, and vegetarian food. Dividing the recipes by these categories would have followed the kind of menu alternatives that the restaurant where he worked offered: one fish dish, one meat dish, and one vegetarian alternative.

John used his log-in name from other online sites. The only personal information he entered was an anonymous email address, a hotmail address. He did not see himself as particularly social in the sense that he immediately goes out and talks to strangers.

John believed that the recommender system selected the recipes that were most popular. Thus, he had both noticed the feature and reflected over it. He found the recommendations good and wanted to try out the recipes. He said that he was always looking out for new recipes and consequently found that the recommender fitted his needs. He did not bother to sort the recipes in any other order such as in alphabetical order. In the actual process of deciding on a recipe, he looked at the ingredients. The other information was of no use to him. As a chef, he could determine the value of a recipe from just looking at the ingredients.

The only functionality John used in the system was to move between collections and look for recipes. He was always visible and did not feel uncomfortable with other users being able to see his actions. John found Kalas useful mainly due to the large number of recipes it contains.

Maria: 68 years old. Maria used Kalas to get recipe hints. She found the system quite useful, but stopped using it when she got a virus-infected mail from another Kalas user.

Maria was open about herself. She used her real name as log-in, partly because it was easier to remember, and partly because she did not mind giving it away. She had no problem giving out her full name, email address, and phone number. She did not see the point in being secret about that despite the subsequent virus-infected email.

Maria was not too comfortable with the way the system ordered recipes. In particular, she did not understand the thumbs-up symbol. She did, however, appreciate and see the point in dividing the recipes into collections that matched the needs of different types of people.

Similar to John, she studied the ingredients when deciding if a recipe was worth cooking or not. She claimed not to have used any other functionality other than just looking at the recipes themselves. However, Maria said that she looked up and read profile data about other people who were logged into the system when she was and actually got in contact with one via email. Clearly, she did not regard those social functions as part of the system functionality. She claimed she would never chat.

The best thing about Kalas was that it offered a new way of finding recipes. The worst thing about Kalas was that it was difficult to install. When shown a picture of a very crowded Kalas with many users in one particular collection, Maria claimed that she would go to the collection that most likely contained the recipes she was after.

Insights from the interviews. Anna, John, and Mia seemed to have some grasp of the social filtering functionality and found it useful. Maria, on the other hand, claimed to not have understood the meaning of the thumbs-up symbol or the recommendations but made use of the other social functions in the system. Anna and Mia actively influenced Kalas by voting and writing comments to recipes. However, only Anna saw those actions as a way to deliberately modify the space.

Anna, Maria and Mia wanted to read about others, chat, and read comments. This shows that they do indeed want to be influenced by what others are doing. Thus, the social layer is a valuable tool in the navigational process.

Giving out personal information is not done automatically but seems to be reflected on from different perspectives, determining when it could provide additional value. Maria read about the other users and contacted them via email, and Anna provided her age as a way to find like-minded users.

6.5 Real-Time Presence, Chat Discussions, and Comments

As indicated previously, there were not enough users logged in at the same time to really create a sense of the real-time presence of others. This, in turn, means that users seldom got the chance to chat with one and other. Out of all the users, only twelve chatted or attempted to chat. There were 2 successful chats between end-users and another 8 successful chats where one of the experimental leaders was involved. On 20 occasions, end-users tried to make contact in the chat, and on another 8 occasions, the experimental leader tried to contact users in the chat. Thus, in total, there were 38 chat occasions. The discussions in the chat did not talk about the recipes but were reflections on the functionality provided by Kalas (with the exception of one single user who talked about having a sausage in her refrigerator). The initiation of one of the two successful chats reveals what kinds of dialogues were attempted:

—Hi Maritha! The program shut down on me before I could answer you. I am 61 years old and live in Norrköping.
—That is a nice city (county). Unfortunately it was a long time since I was there. What do you think about this Kalas-system?
—I have been pretty lonely when I have been in Kalas. Otherwise I find it to be a nice idea. Though I would like another categorization or way of dividing the recipes on for example starters and similar. [. . .]

There were 11 comments on recipes. These comments were tightly tied to the recipes as in, for example, this comment:

> "The plum pudding tasted good and was easy to cook. Maybe it could have been improved somewhat through spreading some potato-starch on the plums. The pie is most surely useful also for other kinds of fruits such as rhubarbs and peaches."

The difference between what is said in the chat versus in recipe comments shows that users understood what sort of social communication was expected in the different parts of Kalas. Our assumption that the topic in chats would be recipes was not correct, at least not with these data. Also, some social communication will simply be for social reasons, not to navigate the space.

7. NAVIGATION AND SUBJECTIVE QUALITY

Given a better feel for how Kalas was used through the statistics, interviews, and description of chat dialogues and recipe comments, we now turn to the effects on navigation and the perceived subjective quality.

7.1 Effects on Navigation

The aggregated or individual traces users leave in Kalas is, as described earlier, supposed to affect the navigation through the system from choosing the recipe collection, choosing among the recipes in the list, and finally, determining whether the particular recipe should be chosen or not. The intended design was that

(1) the real-time presence of users in different recipe collections and how they move between them would influence which recipe collections users move to;
(2) the choice of recipes from the list in each collection would be influenced by the thumbs-up symbol next to the recommended ten recipes shown at the top of the list in each of the nine recipe collections;

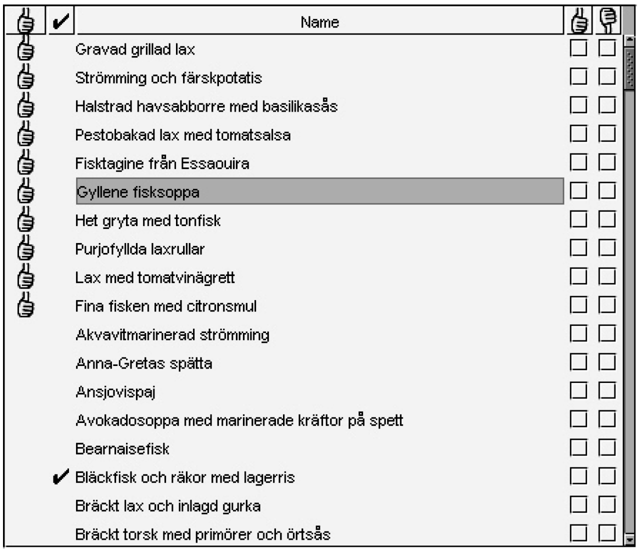

Fig. 4. Kalas recipe list.

(3) the figure indicating the number of downloads of each recipe, the author of the recipe, the comments on a recipe, and possibly discussions about recipes in the chat part of Kalas would help the user to either choose or continue to look for recipes in the collection.

7.1.1 *Moving Between Recipe Collections: Will People Attract People?.* The navigation to find recipes in Kalas starts when the user chooses which recipe collection to go to. We hypothesized that the more users in a collection, the more inclined users would be to go there. Most times, there were not enough users logged into the system at the same time so our analysis here is based on sessions with at least 5 users logged in and sessions with at least 10 users logged in. In the first condition, 27 moves were to empty collections, 61 moves to collections with at least one user, and 23 to the most populated collection. In the latter, 7 moved to empty collections, 36 to populated collections, and 12 to the most populated collection. The pattern is fairly stable in both conditions with slightly more movements to the populated collections than to an empty collection. It should be pointed out that we do not know whether they move to the collection because they see others there or if they would have moved there anyway. On the other hand, the interviews indicate that users would have been influenced by the presence of others even if the statistics do not provide us with any clear evidence.

7.1.2 *Filtering by Recommendations.* The next step, once the user has chosen a recipe collection, is to pick one of the recipes in the scrollable list, or possibly to reorganize it, for example, in alphabetical order. To determine the influence of the thumbs-up symbol next to each of the ten recommended recipes at the top of the list (see Figure 4), the following analysis was done.

Users acted on, that is, saved, printed, voted for, or added to their shopping list, 1129 recipes during the period of use, see Table III. Out of those, 199 (18%) recipes were recommended to them, while 305 (27%) were found through the search functionality. The remaining 625 (53%) recipes were found by scrolling the list beneath the recommended recipes. Thus, the recommender system is used slightly less than explicit search. This is interesting given that recommendations are not explicitly sought and do, therefore, not require any extra actions on behalf of the user as explicit search does. If those 199 recipes had not been among the first ten, the user could only have found them by scrolling to see them further down in the list, or search for them explicitly. Another fact that shows the complementary roles of explicit search and recommendations is that none of the 20 most popular recipes chosen from the recommended recipes and explicit search were the same recipes.

It is known that users are quite influenced by what they can see on the screen and scrolling down to see more recipes requires more energy than just choosing among the ones visible in the window. Thus, since 18 recipes were visible in the window (see Figure 4), it is interesting to know how many of the eight visible, but not recommended recipes, were chosen compared to the ten recommended. It turned out that they accounted for 90 of 1129 selected recipes. Thus, about 1% of the recipes in the last eight positions were chosen, while 1.76% of the recipes in the first ten positions were chosen. It should be noted that users could decide to organize recipes alphabetically, by date, or by chosen recipes, and would therefore not always see the recommended recipes in the first ten positions.

Digging deeper into the usage logs, we noted how users' choice of the first recipe they ever pick in Kalas differs from their consecutive choices (see Figure 5)[2]. The trend seems to be that users will not pick recommended recipes as often the first time compared to their subsequent selections. There are two possible explanations for this result. Users might be learning the meaning of the thumbs-up symbol after some time and only then will they see the point of the first ten recommended recipes. A second explanation might be that the recommendations are somewhat random until the user has selected at least one recipe. It seems more likely that the second explanation does not hold since the recommendations will not substantially change after only one choice. The difference between the third and fourth graphs, depicting choices of second versus subsequent choices, shows that the pattern quickly stabilizes and recommended recipes continue to be chosen to a similar extent.

It could be argued that implicit feedback (such as printing or saving the recipe) accounts for a large number of the votes on recommended recipes and thus the 199 chosen recommended recipes may not accurately reflect that users liked the recommended recipes. If we make a comparison with all implicit votes removed, 31 out of 110 explicit votes were on recommended recipes. That is, approximately 28% of the explicitly valued good recipes were also recommended.

[2]The number of choices of recipes in positions 11 to 13 (in all four graphs) is unusually high which turned out to be due to an extreme user that selected the same recipe multiple times.

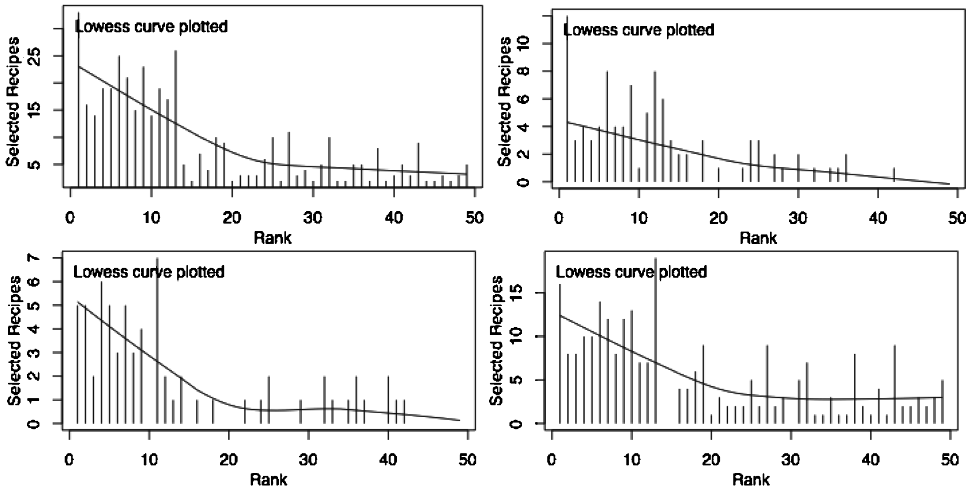

Fig. 5. The four graphs show from which position in the recipe lists people chose their recipes (the first 10 positions are recommended recipes) under various conditions. In the upper left corner, all chosen recipes for the whole period of use, is plotted. In the upper right corner users, first choice of recipe is plotted, in the lower left corner, their second choice of recipe, and in the lower right, their third and subsequent choices are plotted.

In the study, seven users gave in total 11 negative votes. This is a very small number and one might be inclined to draw the conclusion that negative feedback is not needed. However, six of the negative votes were preceded by an implicit positive vote on the same recipe thus negative votes constitutes a way of correcting the system's profile. We would like to argue that negative votes are important in systems that build user profiles from implicit positive votes.

7.1.3 *Choosing or Not Choosing the Recipe.* Once the user is looking at a particular recipe, choosing it or not choosing it will depend on many factors: the ingredients, the explanation of how to cook it, the picture of the dish, and so on. The social texture added in Kalas includes the author of the recipe, a figure describing the number of downloads, and any comments left by previous users. To know whether this social texture influenced users' choice, there is no reliable data to access but their own accounts of whether they believe that it was useful information. A series of questions were asked in the final questionnaire (see Table V that follows) to determine the general attitude towards the chatting possibility, the ability to comment, the ability to vote on recipes, and the ability to see other users. While chatting rarely happened and there were few comments in the system, some users still rated those as potentially important in choosing a recipe: 20 users out of 67 rated comments as an important factor, while only 7 users claimed the recipe author to be an important factor. The author or which organization provided the recipe seems uninteresting.

7.2 Perceived Subjective Quality

Let us move on to whether the trails had any influence on the more subjectively perceived qualities of the system. As discussed previously the benefits of social

Table V. Questionnaire Replies on a Scale from 1 (Negative) to 7 (Positive)

Post-questionnaire Question	1	2	3	4	5	6	7	Median (Mean)
How important is the recipe author in your choice of recipe?	43	7	7	9	3	3	1	1 (2.1)
How important is the recipe source in your choice of recipe?	35	14	5	12	6	1	0	2 (2.2)
How important is it that others have chosen the recipe in your choice of recipe?	35	13	6	13	4	2	0	2 (2.2.)
How important are a recipe's comments in your choice of recipe?	11	12	12	12	12	8	0	3 (3.4)
How important is it that a recipe was recommended in the chat in choosing a recipe?	13	11	12	11	8	7	1	3 (3.2)
What do you think about the possibility to comment recipes?	2	2	2	21	12	16	12	5 (5.0)
What do you think about the possibility to chat?	2	5	4	36	10	3	3	4 (4.1)
What do you think about the possibility to see others?	1	1	2	28	8	4	2	4 (4.3)
What do you think about the possibility to vote on recipes?	2	1	1	23	14	15	5	5 (4.8)
What do you think about the recommended recipes?	0	0	3	29	20	18	2	5 (4.8)

navigation are not only to accumulate users' trails and by those guide others in navigating large information spaces, but also to provide users with a sense of social presence, of not being alone in the space and a subjective stance—a texture of others' judgments that will aid their choice and make them more confident in trusting their choice.

As argued earlier, the result is of less importance from a social navigation perspective unless users in some sense notice and also understand that the trails actually come from other end-users. This is a particularly interesting problem for the recommender functionality.

7.2.1 *Do Users Understand Where Recommendations Come From?*. Did the users understand that it was socially computed recommendation that made some recipes turn up at the top of the list with a thumbs-up symbol? In the final questionnaire, we asked: "How do you think Kalas chooses recipes for you?" We gave them six different alternatives, listed in Figure 6. A total of 34 of the 73 users believed that other users' choices affected the recommendations. As many as 28 users did not even reflect on where the recommendations came from. As can be seen in Table V, in the last question, 29 users were neutral (grade 4 on the 7-grade scale) to the value of the recommended recipes. It might be that these users are either unaware of the meaning of the thumbs-up symbol or have actively decided not to follow the recommendations.

We compared users' own beliefs on the explanation to the thumbs-up symbol with how they actually chose recipes. The 34 users who believed that recommendations were products of other user actions chose 60 recommended recipes of 273 recipe choices, that is, 22% of their chosen recipes were recommended.

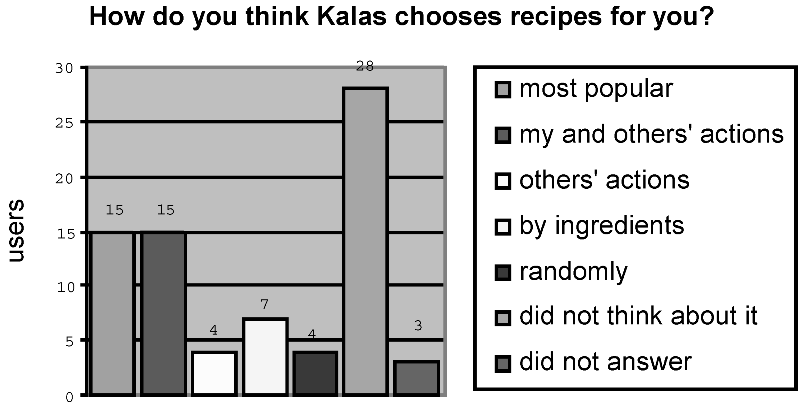

Fig. 6. User replies to the question "How do you think Kalas chooses recipes for you?"

The 28 users that did not reflect on where the recommendations came from chose 32 recommended recipes from a total of 111 recipe choices, that is, 29% of their chosen recipes were recommended. The 11 users who misinterpreted the meaning of recommendations chose 19 recommended recipes from the in total 253 chosen recipes, that is, 8% of their chosen recipes were recommended. Thus, interestingly, the second group selected a large proportion of their recipes from the recommendations perhaps indicating that users who did not reflect on the meaning of the thumbs-up symbol might still have acted on it unconsciously?

The in-depth interviews show that some users do understand the meaning of the thumbs-up symbol and follow it (John and Anna). Mia on the other hand, both voted explicitly and made use of the recommendations even if she did not understand them fully. Finally, Maria had not made use of them nor understood their meaning.

As Kalas did not attempt to explain its functionality, we believe that it is still a quite strong result that about 34 subjects roughly understood where the recommendations came from and that about 40 subjects were positive about the recommended recipes, see Table V. We believe that this is partly due to the thumbs-up symbol in Kalas that seems to provide an intuitive and reasonable understanding of the recommender functionality as it is easily associated with voting. It is also reused for both positive and negative (thumbs-down) votes in the private collection of saved recipes that each user has. When logging onto Kalas, users always start in this personal collection (i.e., My Box, see Figure 7) and are thus confronted with all of their past choices. Next to each such choice, thumbs-up and thumbs-down checkboxes are available.

A difference between Kalas and other recommender systems is that users do not have to make a number of initial selections before receiving recommendations. On the one hand, this could be the explanation as to why so many users did not understand where the recommendations came from or why they could be relevant to them. On the other hand, when users initially have to put in a lot of work to bootstrap their profile, the threshold to start using the system increases. When recommendations are used as a tool that is well integrated

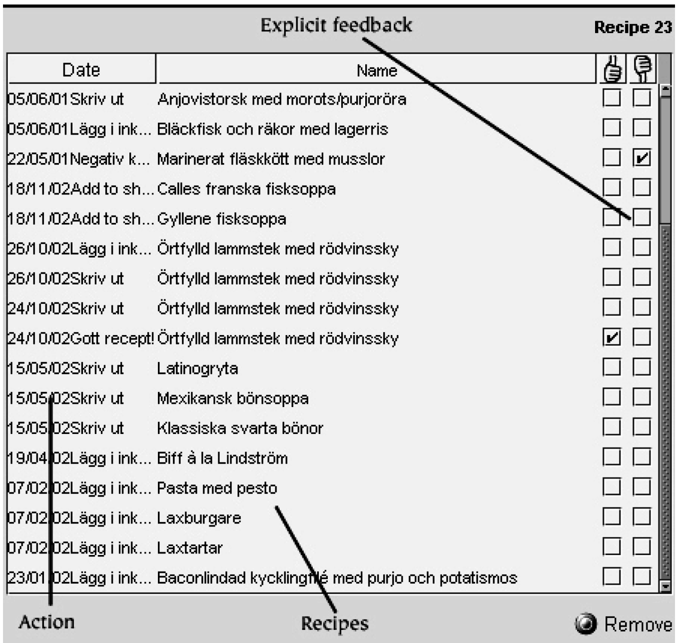

Fig. 7. My box.

with the rest of a system, it is not that important that recommendations are perfect.

7.2.2 *Are They Influenced by Social Trails?.* Table V (the first five questions), implies that the social texture is not that important when choosing a recipe. The more explicit social texture—comments and recommendations in chat—looks to be more important than other social information attached to recipes. In fact, 20 users out of 67 rated comments as an important factor, while only 7 users claimed the recipe author to be an important factor. Within the social information, we find large differences in terms of value. Both comments and chatting are in a way personal and directed to others which could account for this difference.

What is interesting to note from Table V is that, while users are only reluctantly interested in the comments of others or what is chatted about in their recipes choices, they still value those functionalities highly. Could it be that most users are more interested in expressing themselves and seeing the trails of others than in actually allowing those factors to influence their own choices? We would like to argue that these functions made Kalas more pleasurable. They enhanced the experience with the system, although they did not explicitly affect the choice of recipes, at least not on a conscious level.

8. DISCUSSION

Both the prestudy and the Kalas study presented here show that people use some aspects of the implemented social navigation features in their navigational process. They seem to be influenced by the recommended recipes, they tend to go to the most populated part of the space when five users or more are logged in, and they are interested in comments on recipes. Perhaps more important is the overall impression that users appreciated the social functions in Kalas and found they liked the idea of using them perhaps more to express themselves than to actually see the comments and choices made by others. We might speculate that they acted more on some of the social trails than they might have understood themselves.

Some features did not work as expected: (1) the chatting facility never took off, (2) users said that they liked the comments and the idea of writing them themselves but did not do so to any great extent, and (3) there were too few subjects most of the time to see any interesting effects of the real-time presence of others. Still, 67 users understood that it was possible to chat, even if chatting was only attempted by twelve users on 38 different occasions. Similarly, 70 of our users understood that it was possible to comment on recipes and they liked the idea but only 11 comments were actually entered in Kalas.

The thumbs-up symbol worked fairly well, 63 subjects understood that they could vote on recipes, but the connection to the actual recommender functionality was only understood by about half the users. They, on the other hand, seemed to be more appreciative of the recommended recipes than those users who had not really noticed them or acted upon them. It might be that a really simple solution such as adding a pop-up help-text when the user moves the mouse over the thumbs-up symbol in the list of recipes could help users to grasp this functionality.

8.1 Bootstrapping Recommendations

Bootstrapping is considered to be one of the more difficult problems to overcome in collaborative filtering. Not until enough users have rated enough items in the database will the recommender be able to predict users' rating with any accuracy. What was interesting in the study of Kalas was that users did not pick more nor fewer recommended recipes after their initial first choice of recipe. More or less from the start, they picked recommended recipes and went on doing so throughout the whole period. It is also notable that 40 (of 73) users liked the recommender functionality (29 were neutral and 3 slightly negative).

Thus, one might speculate that bootstrapping is not necessarily such a big problem if we look at a recommender system from a more utility-based perspective. In domains where many items fit well with users' needs and the items are all of fairly high quality so that users will rarely be disappointed, it might be enough that the recommender starts with showing the overall most popular items. As the recommender improves its performance, the individualized recommendations will be highly useful to users but getting the most popular to start with is good enough. Cosley et al. [2003] show that users grades of movies, for example, are not very stable; they can be influenced to change their ratings based on what the system predicts for them. As discussed previously, the perceived quality of a recipe might be higher if it is recommended and if the system does not disappoint the user by recommending a bad recipe. Finding the best one might not be absolutely necessary nor always possible. The benefits of not having to bootstrap the system is mainly to the end-users: they do not have to rate a lot of items in the system before they can get recommendations as is the case with many recommenders.

8.2 Evaluating Recommendations

The previous point inevitably leads to the next point, namely, how to evaluate a recommender system. A traditional analysis of the recommender would not have shown that bootstrapping is less of problem than expected. Looking at pure algorithmic properties, the system (e.g., precision and recall) will not accurately measure the performance of a recommender system. Instead, we have to focus on the utility in terms of how and why users act on recommendations. Both the subjective views and usage analysis suggest that people like recommendations. As with the other social features, recommendations also inspire people and help them explore the space. Precision and recall will not evaluate the entire benefit of including a recommender system in the system nor gauge how well the domain fits with the recommender functionality.

8.3 Problems in Evaluating Social Navigation

The results showing that users pick recommended recipes despite not having reflected on the meaning of the thumbs-up symbol, leads us to speculate that a lot of the social trails are created and used without users consciously noticing them. Even when the social features are noticed and understood, many users claim that they do not necessarily feel that they support their navigation in terms of getting from one location to another, again despite the fact that the statistics seem to say otherwise. Social affordance and following users around might be activities that take place without users realizing it, or perhaps it is not really seen as a vital part of the decision process. For example, if asked why you go to a particular restaurant, you might say that it is because of the food but you will probably not say that it is because of the number of people who frequent it.

Therefore, we believe that social navigation should not be evaluated in the same way as more traditional navigational aids such as maps or signs (that have a very clear purpose). We have seen that social quality is not used in the same direct way in the decision process as, for example, recipe ingredients. The four in-depth interviews showed that users separate recipe-searching from the other social activities. Three of four in-depth interviewees stated that it was the ingredients that made them choose a recipe. This is why we choose to not only see social navigation as aiding in moving most efficiently from point A to point B, but also adding to perceived quality and, in general, adding to the whole experience of using the system. Thus, many different sources of data from users must be collected besides the more objective log statistics. In those data, it must be possible to separate the reasons why a particular item or route through the system is selected—beside the content, how are users affected by the social texture? Some of this was captured in our questionnaires and interviews but this could be improved.

8.4 Finding Enough Users

While we were proud to have so many users in a system designed and implemented by researchers (and in fact, still have users who logged in even after the study was completed), we did not have enough users to evaluate all aspects of Kalas. The real-time presence, chatting, and comments on recipes needed more subjects for a complete evaluation of the system. Because we believed that it was important to find a user community who were already interested in food recipes, and we did not recruit subjects among friends and colleagues, we were basically restricted to the ones that came through the hemma.net site.

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers who gave unusually insightful and helpful comments to the first version of this article. We also wish to thank the all of the anonymous users who took part in the study.

REFERENCES

BREESE, J. S., HECKERMAN, D., AND KADIE, C. 1998. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence*. 43–52.

CALDERÓN-BANAVIDES, M. L., GONZÁLEZ-CARO, C. N., PÉREZ-ALCÁZAR, J. DE J., GARCIA-DIAZ, J. C., AND DELGADO, J. 2004. A comparison of several predictive algorithms for collaborative filtering on multi-valued ratings. In *Proceedings of the ACM Symposium on Applied Computing*. 1033–1039.

COSLEY, D., LAM, S. K., ALBERT, I., KONSTAN, J., AND RIEDL, J. 2003. Is Seeing Believing? How Recommender Systems Influence Users' Opinions. In *Proceedings of the Conference on Human Factors in Computing Systems*. Fort Lauderdale, FL. 585–592.

COSLEY, S. LAWRENCE, AND D. M. PENNOCK. 2002. REFEREE: An open framework for practical testing of recommender systems using researchindex. In *the 28th International Conference on Very Large Databases (VLDB'02)* Hong Kong.

DIEBERGER, A. 1997. Supporting social navigation on the World-Wide Web. *Intl. J. Hum.-Comput. Stud. 46*, 6, 805–826.

DIEBERGER, A., DOURISH, P., HÖÖK, K., RESNICK, P., AND WEXELBLAT, A. 2000. Social navigation: techniques for building more usable systems. *Interactions* (Nov.-Dec.), ACM.

DIEBERGER, A. AND LÖNNQVIST, P. 2000. Visualizing interaction history on a collaborative Web server. *Hypertext 2000*. 220–221.

DOURISH, P. AND CHALMERS, M. 1994. Running out of space: Models of information navigation. In *Proceedings of Human Computer Interaction (HCI'94)*.

ERICKSON, T. 1996. The World Wide Web as social hypertext. *Comm. ACM 39*, 1, 15–17.

ERICKSON, T. 2004. Designing online collaborative environments: Social visualizations as shared resources. In *Proceedings of the 9th International Working Conference on the Language-Action Perspective on Communication Modeling (LAP'04)*. M. Aakhus and M. Lind, Eds. Rutgers University, The State University of New Jersey, New Brunswick, NJ.

ERICKSON, T., SMITH, D. N., KELLOGG, A. W., LAFF, M., RICHARDS, J. T., JOHN T., AND BRADNER, E. 1999. Socially translucent systems: Social proxies, persistent conversation, and the design of "babble". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Pittsburgh, PA. 72–79.

GIRGENSOHN, A. AND LEE, A. 2002. Making Web sites be places for social interaction. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW'02)*. 136–145.

HERLOCKER, J., KONSTAN, J., AND RIEDL, J. 2000. Explaining collaborative filtering recommendations. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. 2–6.

HILL, W. C., HOLLAN, J. D., WROBLEWSKI, D., AND MCCANDLESS, T. 1992. Edit wear and read wear. In *Proceedings the ACM Conference on Human Factors in Computing Systems*. Monterey, CA. 3–9.

HÖÖK, K., MUNRO, A., AND BENYON, D., EDS. 2002. *Designing Information Spaces: The Social Navigation Approach*. Springer Verlag.

KONSTAN, J. A., MILLER, B. N., MALTZ, D., HERLOCKER, J. L., GORDON, L. R., AND RIEDL, J. 1997. GroupLens: Applying collaborative filtering to usenet news. *Comm. ACM 40*, 3, 77–87.

MAGLIO, P. AND BARRETT, R. 1999. WebPlaces: Adding people to the web. In *the 8th International World Wide Web Conference (WWW8)*. Toronto, Canada.

MUNRO, A. 1999. Fringe Benefits: An ethnographic study of social navigation at the Edinburgh festival. Deliverable 2.1.1 from the PERSONA Project, available from SICS, Stockholm, Sweden.

MUNRO, A., HÖÖK, K., AND BENYON, D., EDS. 1999. *Social Navigation in Information Space*. Springer Verlag.

SHARDANAND, U. AND MAES, P. 1995. Social information filtering: Algorithms for automating "word of mouth". In *Proceedings of Human Factors in Computing Systems (CHI'95)*. Denver, CO. 210–217.

STREETER, L. A., VITELLO, D., AND WONSIEWICZ, S. A. 1985. How to tell people where to go: Comparing navigational aids. *Intl. J. Man-Machine Stud. 22*, 549–562.

SVENSSON M., HÖÖK, K., LAAKSOLAHTI, J., AND WAERN A. 2001. Social navigation of food recipes. In *Proceedings of of the Human Factors in Computing Systems (CHI'01)*. Seattle, WA. 341–348.

TIMPKA, T. AND HALLBERG N. 1996. Talking at work—professional advice-seeking at primary healthcare centers. *Scand. J. Prim Health Care 14*, 130–135.

WEXELBLAT, A. 2002. Results from the Footprints project. In *Designing Information Spaces: The Social Navigation Approach*, K. Höök, A. Munro, and D. Benyon, Eds. Springer Verlag.

WEXELBLAT, A. AND MAES, P. 1999. Footprints: History–rich tools for information foraging. In *Proceeding of the Conference on Human Factors in Computing Systems (CHI'99).* Pittsburgh, PA. 270–277.

Received September 2003; revised July 2004; accepted March 2005 by John Riedl and Paul Dourish