‘Knowing me, knowing you’ — using profiles and social networking to improve recommender systems

P Bonhard and M A Sasse

The Internet has not only brought us more information and choice, but has also increased the burden of making a choice. Recommender systems aim to address this problem by providing personalised recommendations in areas such as music, films or books. Research on recommender systems has focused on improving the matching algorithms. The research presented in this paper takes a user-centred approach. Since recommendations are usually presented as lists of items, without explanations or justifications, users struggle to find out how appropriate recommendations are for them. Our research has shown that the relationship between advice-seeker and recommender is extremely important, so ways of indicating social closeness and taste overlap are required. We thus suggest that drawing on similarity and familiarity between the user and the persons who have rated the items can aid judgement and decision making. This was tested in an experiment, which carefully controlled familiarity, profile similarity and rating-overlap between the user and those rating items. The results help us understand the decision-making processes in an on-line context, and form the basis of a user-centred recommender system approach. We suggest that recommender systems can be improved by combining the benefits of social networking applications — such as explicit networks of trust — with the matching capabilities of recommender systems.

1. Introduction

Internet users today are confronted with an abundance of information. Whether one is comparing prices for CDs and DVDs, searching for film reviews, or looking for a restaurant for an important date, the available information at hand exceeds the amount of information most users want to consider before making a choice. In effect, choice has become a burden [1].

Recommender systems (RSs) have been developed to help people deal with this problem by filtering the available options according to their preferences. Amazon, for instance, recommends items based on our purchase and search history, while MovieLens [2] lets us rate films and generates recommendations based on those ratings. At the heart of each RS is a large database of:

- items (such as books, films, etc),
- users,
- their ratings of those items.

Recommendations are derived by calculating similarity of ratings across different users. Currently, RS provide the user with a ranked list of items, based on the ratings overlap with other users. They provide little or no information about where these choices come from, and why some items have been prioritised over others.

These systems aim to emulate the process of seeking advice from trusted sources, yet users still struggle to convert the data they receive from such systems into meaningful information. When people seek advice from peers or trusted sources in the real world, there are a variety of interpersonal cues, an interaction history, and opportunities to question the advice given, which are missing in an on-line system. Therefore when users receive recommendations from an RS without any meaningful explanation, the recommendation lacks the cues associated with one from a real person and appears more like a search result.

This is a problem an RS should solve rather than create. Similarly, since people inherently know which of their friends to trust for a particular recommendation, RS should imitate this naturally occurring evaluation and decision-making process. It is therefore necessary to understand how people arrive at their decisions in such domains and which recommender characteristics are important to them.

Human-computer interaction (HCI) approaches to RS research have primarily focused on evaluating the user interface of existing RSs [3—5].

Our research, however, took a step back and asked when users seek recommendations, and what information they hope to receive. We translated our findings into a mock-up design of an RS which incorporates social networking features, and then evaluated those designs through experiments. This paper first presents a brief review of RS research to date, followed by an overview of the psychology literature on advice-seeking and decision making. We then consider social networking applications and how they fit into the RS context. This is followed by a section presenting the qualitative and quantitative studies exploring the integration of RS and social networking. The last section then discusses the implications of this research for RS design and future research.

2. Background

RSs have been deployed in various eCommerce contexts, such as book or music shopping, e.g. Amazon (www.amazon.com), general rating sites (www. ratingzone.com), and specific rating sites such as MovieLens (http://movielens.umn.edu).

RSs aggregate the information received and redirect it to the appropriate recipients. While many strategies for computing recommendations have been explored, such as item-based collaborative filtering [6], Bayesian networks [7] and factor analysis [8], user–user collaborative filtering (CF) comes closest to emulating real-world recommendations because they are based on the user rather than item matching. Recommendations are generated for a given user by comparing their existing ratings to those of all other users in the database. In doing so, a neighbourhood of similar users is established, and based on that, rating predictions are computed for items that users have not yet rated, but closest neighbours have.

2.1 Recommender systems research

RS research to date has focused on designing algorithms for more effective and efficient computation of rating predictions. The former aims to increase the precision of predicting ratings. This is tested through existing rating data sets, where part of the rating set is deleted, and the prediction results from algorithms are compared against the real ratings. Prediction efficiency is concerned with the computational cost in terms of time and resources for calculating these predictions.

HCI approaches to RS research have been limited to examining existing RSs in order to establish interaction design guidelines [3—5].

When comparing how advice from friends and from RS is perceived, Swearingen and Sinha [4, 9] found that, while people overall preferred recommendations from their friends, they appreciated the ability of an RS to provide serendipitous recommendations that broadened their horizons. In that context, Swearingen and Sinha [5, 9] identified two factors as fundamentally important in the overall usefulness of an RS — familiar recommendations and system transparency.

Familiar recommendations can come in different forms — items previously consumed, or items that are related to known items (e.g. books by the same author). Familiar items can generate trust in the system, but it can also make the recommendations seem too simplistic. As for transparency, understanding the inference leading to a recommendation (and agreeing with it) not only increases trust in the recommendation, and the system providing it, but also makes it more likely that the user will follow the recommendation.

Using a similar approach, Herlocker et al [3] conducted an extensive study examining what effect explanations for collaborative filtering results have on the user's perception of the system. In testing different explanation interfaces, they found that explanations are important to users, because their own reasoning often does not match the inference mechanism of the system. Users were less likely to trust recommendations when they did not understand why certain items were recommended to them. Herlocker et al [3] suggest that a rating histogram of the user's closest neighbours is the most effective way of explaining the results of collaborative filtering.

These above studies took an evaluation approach to improving an existing RS. In our view, however, the user requirements for RSs have not been effectively investigated. Rather than continuing the current cycle of deploying a matching algorithm and seeing how users respond, we took a step back and examined existing literature on advice-seeking and decision-making strategies, to identify what support users seek during these activities.

2.2 Trust research and recommender systems

Previous HCI studies on RS aimed to increase user trust in these systems by helping users understand how the system calculates recommendations. Related HCI research in recent years has investigated trust in eCommerce systems [10], social networking, virtual communities and recommender systems [11, 12]. Each of these aims to address different aspects of the concept of trust depending on the definition applied.

Massa and Avesani [11] addressed the problem of data sparseness in collaborative filtering — in a huge database, on average two users are unlikely to have rated the same items in a similar fashion. To overcome this problem, they create a trust graph based on the degree of connectedness of users, using the distance in terms of arcs between two users on this graph as a measure of trust. This transitive trust graph is then used to increase the number of comparable users.

O'Donovan and Smyth [12] take a different approach by using past rating reliability to generate trust values that increase or decrease the weight given to predictions of neighbouring users. They divide trust into two possible categories — profile level and item level trust. On the profile level they compare all common existing ratings of an advice-seeker and a recommender, and examine whether the recommender would have been able to predict the correct ratings for a given advice-seeker. The percentage of correct predictions across all common items is their profile-level trust value. Item-level trust is more fine-grained and only measures the percentages of recommendations for one item that were correct.

Each of these approaches addresses a different challenge for collaborative filtering (CF) algorithms. Whereas the Massa and Avesani [11] approach can be used to overcome a technical problem in order to generate recommendations more quickly, O'Donovan's [12] mainly aims to increase accuracy (precision) of such a CF algorithm. However, both approaches address a particular technical problem without considering the user's perspective on trust.

Riegelsberger et al [10] developed a trust model for interactions between two actors in a situation of uncertainty and risk, such as an exchange of goods, information or advice. The model divides means of signalling trustworthiness into symbols and symptoms, both of which the technology involved can transmit. Symbols can be eCommerce trust seals or a professional looking Web site. Symptoms, on the other hand, cannot be specifically created; rather they are a by-product of trustworthy actions. In the context of an RS, a recommender who repeatedly gives good advice to an individual shows evidence of trustworthiness. Someone who has given good advice to others may become a symbol of trustworthiness.

Briggs et al [13] identified trust-warranting factors for on-line advice. They concluded that three factors were paramount for users to perceive information as trustworthy:

- source credibility,
- advice personalisation,
- predictability.

Source credibility refers to the overall impression of a site, personalisation to whether the information is tailored to the user's needs, and predictability to whether the information presented reflects the user's knowledge, and prior experience with this and other sites. However, Briggs et al actually examine information seeking, rather than advice-seeking, and which elements of a site design make the information appear trustworthy.

In RS research, there has been a significant lack of considering established decision-making and advice-seeking research. For an RS to address users' needs in this domain, however, the requirements of the user's task need to be considered.

The next section gives a brief overview of the relevant psychology literature on this issue, and identifies the support requirements which can be inferred from it. A more detailed description can be found in an earlier paper [14].

2.3 Decision making and advice seeking

The main purpose of an RS is to aid its users in their decision making by presenting a reduced set of options in the form of advice or recommendations. To understand how people normally deal with advice from other sources, it is instructive to examine the decision-making and advice-seeking literature in psychology.

2.3.1 Decision making

Classic decision-making theory proposes various strategies that decision makers might adopt when examining options [15, 16]. Utility theory, which falls under the compensatory models of decision making [16], would predict that people make decisions based on the expected utility of an option. Evidence, however, suggests that decision makers often use various, less precise non-compensatory models. Decision makers often do not make decisions in the rational or logical way proposed by utility theory, but rather they are often influenced by a plethora of non-rational factors.

2.3.2 Advice seeking

The advice-seeking literature has predominantly focused on objective domains where advice can be classified as right or wrong [17—20]. Generally, advice seeking is seen as a problem of combining and weighting different information sources to come to a final conclusion.

Whereas the benefit of combining opinions in matters of fact is both demonstrable and understood theoretically, the same is not true for matters of taste [18]. Simple aggregation of opinions in matters of taste raises conceptual difficulties as people are entitled to different opinions on movies, music or restaurants. Yaniv argues that the most promising strategy to examine these factors in matters of taste would be in considering the personal match between advice-seeker and recommender, assuming that the greater the similarity between them, the greater the impact and benefit of advice received [18]. The argument is therefore not to find optimal solutions, but rather to use heuristics that simply find good solutions.

2.3.3 Social embedding of recommendations

Researchers have pointed out that recommendations have to be considered in the social context in which they are delivered [21]. For instance, what kind of occasion is a restaurant recommended for? Is it a business dinner or a first date? Who are you going to the cinema with? Are you buying a CD for yourself or as a present?

RSs operate under the assumption that the world can be described objectively, and that optimal solutions to problems can be deduced from these objective descriptions [22]. Consequently, recommendations from such systems are de-contextualised, thus ignoring the situatedness of a recommendation.

To allow advice-seekers to judge the validity and appropriateness of a recommendation, situatedness is a crucial piece of information that should be included in an RS.

Situating recommendations in their social context in an on-line environment leads us to examine how people use tools to socialise on line. Social networking applications provide a platform for their users to connect their friends through their profiles. Seeking recommendations from friends is a naturally occurring social process, which has so far been under explored in this context. Section 2.4 therefore briefly examines social networking applications.

2.4 Social networking

Social networking applications such as Friendster [23], Facebook [24], hi5 [25], and mySpace [26], have grown significantly over the past few years. While early systems saw an initially enthusiastic take-up, this was followed by an equally quick decline, which was mainly due to a lack of clearly defined usage goals. Once signed up, users simply did not know what to do with the applications. Donath and Boyd point out [27], these networking sites had the three following basic underlying assumptions:

- that there is a need for people to create connections,
- that using a network of existing connections is the best way of doing this,
- that making the above easy is a great benefit.

While users have various options to find other users with whom to make connections, the main one is still through the networks of their friends. Other options usually include some form of database search where interests and demographics can be specified. In networks of potentially thousands of people both these options can be quite cumbersome and do not actually follow the advantages of real-world introductions.

In real-world introductions, the person introducing two people usually has a motive for doing so, be it for professional or personal reasons. In an on-line social network, it is easy for users to browse and contact people in their friends' networks, thus skipping the introduction stage. Further, in the real world, social connections are typically revelatory about the social status of a person, and seeing someone within the context of their connections is a useful source of information about that person [27].

In an on-line context such as a social networking application, the situation is slightly different. Whereas real-world connections carry some significance (because both parties are tied to their real-world identities and their connection involves some common ground, whether personal or professional), the same is not true for on-line social networking. On-line identities can often be easily altered and thus do not carry the same reliability factor that real-world identities would.

Thus, an often seen phenomenon in on-line social networking is that a lot of users sign up and simply grow their networks as much as possible. Subsequently, their network loses value in terms of situating that person within a social circle.

Both the systems and its users have matured in terms of usage and specific usage goals. Through higher broadband Internet proliferation, more users have access to effectively use these communication platforms for a variety of purposes such as dating, business networking and sharing of digital content.

While certain social networking platforms are still completely open (e.g. Friendster), some opt for a semi-closed approach (e.g. Facebook), and others for a closed 'by-invitation-only' approach (e.g. Orkut). Open networks such as Friendster are able to attract a large number of users, but are also more vulnerable to misuse.

Whereas Friendster lets users sign up without knowing anybody within the system, Orkut only allows users access if they have been invited by an existing member. The logic of this 'by-invitation-only' strategy is to have a trusted network of identifiable contacts and thus to stop people from abusing or attacking the system.

Facebook offers a different approach in being a semi-closed network that only university students/staff/alumni can sign up to (verified through their university e-mail extensions, e.g. ucl.ac.uk or mit.edu). This semi-closed approach carries the advantage that because the user accounts are tied to university e-mail addresses, it is harder for users to switch identities by signing up with a different e-mail account.

In addition to that, as a default Facebook limits profile viewing to people within either the same university or the same geographical location. In this way, users are primarily encouraged to seek contact within their existing and potential real-life social network based around their university. This forcefully bridges the gap between people's real life and on-line networks in that there is an increased likelihood that two connected people actually know each other through university.

Users' profiles in these networks are not only about depicting the users themselves, but also a way of communication, which helps to '... shape the representation of others within the system' [28]. The context in which this communication takes place, however, is very particular because (apart from their immediate friends) users need to negotiate their profile's representation to an often largely unknown audience. Whereas in real life, people have clear distinctions among their public, professional and private lives, in an on-line context this distinction is lost and thus 'people must be prepared to explain both their performance and that of their friends' [28].

Social networks are a new form of self-representation and communication, and subject to a social behaviour that is different to the real world. For instance, while for some users the main purpose of social networking sites has been to amass social currency through a huge network of friends (a large number of whom they do not actually know in person), for others there is a more goal-driven approach (e.g. business in 'LinkedIn'), where accountability is valued more highly, albeit at the cost of reduced privacy because people use their real names [27].

As the main point of social networking sites is making new connections, the underlying assumption is that having a mutual acquaintance or being connected via a chain of acquaintances, provides context for connecting. People need a focus to rally around to turn an encounter into a connection, for which there must be some common ground. Finding a mutual acquaintance establishes a common ground. This includes beliefs and interests, which through the mutual acquaintance are assumed to be shared to a certain extent [27].

Apart from mutual acquaintances, however, finding like-minded people is still a very laborious search process in which the user has to browse different networks, communities and profiles, and make individual judgements about whether another person is actually of any interest.

2.5 Finding common ground — social recommender systems

While ratings are the essential ingredient of any RS, obtaining them is often difficult, for a number of reasons:

- explicitly rating items or adding items to a database can be a cumbersome process,
- some users quickly need a recommendation as a one-off, and do not see any reason for interacting with such a system over a long period of time,
- having an RS merely for information retrieval when needed is not going to motivate users to either contribute to, or use the system, in the long run.

There is more to advice-seeking and decision making than merely being presented with options. This information needs to be qualified on different levels, such as:

- how the recommendation was computed,
- if done by matching, who it is based on,
- information about the recommended item itself, of course.

Thus, the usefulness of such systems needs to extend that of mere information retrieval, or has to be embedded into another structure that complements the uses of it.

Advice-seeking research has so far focused on objective, factual domains as opposed to domains of taste, with which personalised RSs are predominantly concerned. Simple aggregation of opinions for the purpose of decision making might help as an indicator of popularity, but does not contain the same statistical benefit of improving judgement accuracy as in factual domains. Indeed Yaniv [18] points out that a theory in matters of taste is needed. Additionally, it needs to be examined whether consulting others' opinions actually helps decision quality (assuming a more precise definition of quality) in reducing some of the inherent weaknesses; these include the decision maker’s failure to generate sufficient alternatives for choice and the tendency to try to confirm rather than challenge prior views. Yaniv suggests that the best way to study this domain is to examine the role of a personal match between the advice giver and seeker. This assumes that the greater the similarity between them, the greater the benefit from receiving advice.

This is exactly what collaborative filtering RSs do — matching users according to their tastes as expressed through item ratings, and basing their recommendations on that similarity correlation.

When meeting new people, one the most widespread forms of getting to know each other is establishing some form of common ground in our likes and dislikes. This is a trust-building exercise as it helps evaluate our counterpart’s personality. Indeed, social psychology has shown that people like others who are, among other things, familiar, similar to themselves and with whom they have a history of interaction [29].

Friends from whom we seek recommendations are not just a source of information for us — we know their tastes, views and they provide not only recommendations, but also justification and explanations for them. If in doubt we can always question their recommendations by simply asking about their reasoning and referring back to previous recommendations we might have received from them.

Seeking and receiving a recommendation is a social activity that often involves the discussion of a particular item. Why did the recommender like it? Would the recommender want to experience/buy it again? Will the experience change after a while?

Our main goal was therefore to explore in more detail the factors that drive people’s decision-making and advice-seeking and how this can be applied to improve RS design.

3. Empirical studies

The studies presented in this paper aimed to address the above questions from different perspectives. The first study was to lay the basis in establishing how people actually seek advice in taste domains. The results would then inform the design of the second study, which aimed to test the findings from the first study in an RS context.

3.1 Study 1 — looking for a good restaurant: who to ask

Since RSs aim to give advice in taste domains, we decided to investigate how people choose what books or CDs to buy, which films to see, in which restaurant to eat or even how to find a reliable dentist.

3.1.1 Method

Twelve semi-structured interviews and five focus groups were conducted with a total number of 44 participants with an age range from 21 to 46, including university students as well as professionals. The aim was to elicit concepts and priorities that are important to decision makers when seeking advice, which would then guide the subsequent focus groups.

3.1.2 Results

To analyse the data and model the advice-seeking strategies and processes, we drew on an established approach from social psychology called grounded theory [30]. We present the key findings here, summarised in Fig 1, while readers can refer to our previous paper for a more detailed account [14].

3.1.3 Domains of interest

Participants clearly differentiated between objective and taste domains. Objective item domains items are characterised by measurable and comparable specifications. These are perceived as being neutral to different ideas of taste. Items that fall into this category are things like electronic goods, computer hardware and software, and cars. Items in taste domains are seen as far harder to find advice on. One item may be rated differently by two recommenders, yet both can justify their ratings. Examples for such items are music, books, films and restaurants.

3.1.4 Item considerations — risk

Regardless of the item domain, a key consideration is the potential risk and the consequences in making a choice. At its most basic and common level, one of the first considerations is a financial one (‘It also comes down to the value of the thing’). Participants generally said the greater the financial risk involved in a particular choice, the more detailed and careful their research would be before coming to a decision.

Risk is associated with the following factors:

- bought goods (e.g. books/CDs), that can easily be returned, bear a lower risk than experienced/consumed goods (e.g. cinema/restaurant visits),
- when another party is involved the risk is higher (e.g. romantic dinner date),
- for services (e.g. lawyers, plumbers, dentists), people look for reassurance in a recommendation.

3.1.5 Recommender considerations — known versus unknown

Across all interviews and focus groups, the relationship with the recommender was seen to be extremely

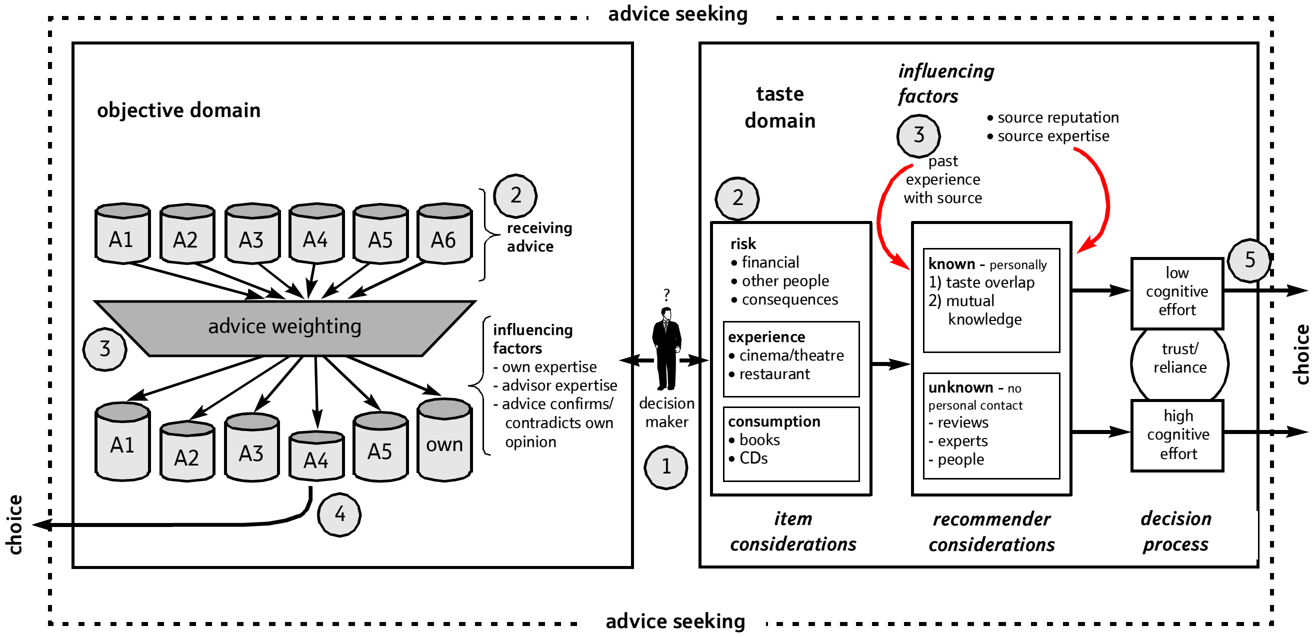

Fig 1 Advice-seeking model.

important when seeking advice in taste domains. While it might seem common sense that people would consult their friends for recommendations for CDs or films, participants clearly pointed out that the relation to the recommender alone is not sufficient. In addition to knowing the recommender, one of two important conditions has to be fulfilled before an advice-seeker will trust a recommender:

- either the advice seeker knows that the recommender has similar or the same tastes (taste overlap),
- or both the advice-seeker and recommender have sufficient mutual knowledge about each others' tastes, so that even with taste differences, the recommender will be able to predict what the advice seeker will like.

3.1.6 Decision process — trust and reliance

Past experience, source reputation and expertise have a significant impact on the final judgement of a recommendation, both from a known or unknown source. They tend to increase or decrease the level of trust in, or reliance on, any given advice. In this context, we define trust as faith in a known advisor in a first time context, whereas reliance is based on past experience. Past experience simply means once advice seekers have received good recommendations, they tend to stick with a particular recommender. Both the reputation and the expertise of a recommender can increase the trust in a first-time encounter. Equally, even if they have not received any advice from a particular advisor in the past, but know that this source is either very knowledgeable or is known to give good advice, this can increase trust in a first-time encounter.

3.2 Study 2 — profile similarity

With the results from the qualitative study in mind, we wanted to examine what effects different recommender characteristics would have on people's choices in an RS simulation. More specifically, what combination of familiarity, profile similarity and rating overlap would have an influence on the choices people make in an RS context? Would a visualisation of profile similarity between the decision maker and recommender influence the decision maker's choice as suggested in social psychology and our previous study [14]? Following Perugini et al's idea of modelling the user [31] and representing their preferences and interests, we aimed to visualise a recommender in a way that would help the decision maker judge the appropriateness of a recommendation. We present an overview of the experiment here while a more detailed account can be found in an earlier paper [32].

3.2.1 Method

Since every participant would be different in terms of demographic data, interests and tastes, we had to create an experiment that would adapt to each individual participant, while conceptually remaining consistent for everyone. To do this, we devised a film festival scenario where participants receive fictitious movie recommendations from recommenders (generated on the fly) that were familiar or unfamiliar, similar or dissimilar, and either had the same or different film tastes (in terms of rating overlap). The individual characteristics of the recommenders were adapted on the fly to each participant.

3.2.2 Participants and procedure

A total of 100 participants completed the study, which lasted 30—45 minutes, in a computer laboratory. The age range varied from 18—44, with a variety of backgrounds, including students and professionals.

Each participant encountered four phases (see Fig 2 for an overview). In the first phase, each participant was required to provide a basic profile, consisting of demographic data such as age, gender and profession, but also preferences and interests such as preferred film genres, hobbies, leisure activities and music tastes. In phase 2, participants rated 20—30 films, which would serve as a basis for generating their recommendations.

In phase 3, participants chose from a series of 48 pairs of films recommended by people who — the scenario made out — had already seen those films.

Phase 4 consisted of a post-study questionnaire, where participants rated some of the profiles they had previously seen in terms of familiarity, profile similarity, rating overlap and trust, and provided qualitative data (in text fields) about their decision-reasoning.

3.2.3 Independent variables — recommender profile characteristics

The recommender profile characteristic variables are listed below.

- Familiarity

  Considering that people consult known sources for recommendations, we aimed to simulate familiarity with a recommender through repeated exposure as suggested in social psychology [29] and the advice-seeking and decision-making literature [33]. Exposing participants to a limited number of profiles before actually receiving recommendations simulates the process of getting to know that particular profile in relation to one's own. This variable had two levels, familiar and unfamiliar. Thus during the rating phase of the experiment (phase 2, see Fig 2), after each rating, participants would see one of four 'buddy' profiles and be told whether these buddies had rated those films the same way or not.

- Profile similarity

  Profile similarity was based on the demographic data, film genre preferences and interests and had two levels, similar or dissimilar. Thus a similar profile would be the same gender, similar age (± two years), same profession and have a significant overlap in their film genre preferences, hobbies and interests. Profile similarity was visualised through highlighting of the interests or preferences the participant and recommender had in common.

  A dissimilar profile, on the other hand, would show little or no overlap in terms of interests and preferences and significantly differ in terms of demographic data.

- Rating Overlap

  Rating overlap was based on the film ratings participants had previously supplied (phase 2). Thus recommenders could either have high or low rating overlap, which was visualised in one of two ways (consistent for each participant):

  — text visualisation (similar to the profile similarity visualisation) showed by explicit highlighting which films the participant and the recommender had rated in a similar fashion,

  — symbol visualisation showed a Venn diagram with two overlapping circles, a large middle circle representing a large rating overlap and a small middle circle representing a small rating overlap.

- Film recommender profiles

  With these three independent variables, each with two levels (e.g. familiar versus unfamiliar), this experiment was a $2 \times 2 \times 2$ design. All possible combinations of these resulted in eight recommender profiles as shown in Fig 3.

3.2.4 Dependent variable — choice

In phase 3 (see Fig 2 for details), participants saw the titles of completely fictitious films and the profile information of the recommender. The reason for not

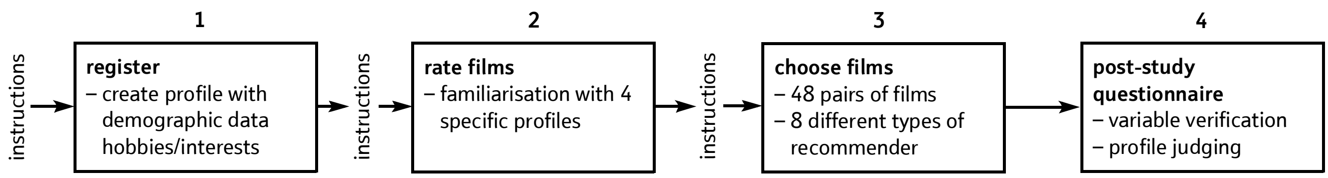

Fig 2 Experiment overview.

profile 1 – familiar – similar profile – high rating overlap	**profile 2** – unfamiliar – similar profile – high rating overlap	**profile 3** – familiar – similar profile – low rating overlap	**profile 4** – unfamiliar – similar profile – low rating overlap
profile 5 – familiar – dissimilar profile – high rating overlap	**profile 6** – unfamiliar – dissimilar profile – high rating overlap	**profile 7** – familiar – dissimilar profile – low rating overlap	**profile 8** – unfamiliar – dissimilar profile – low rating overlap

Fig 3 Recommender profiles.

providing any fictitious reviews or synopses was to minimise noise through a possible participant bias towards any film based on the film's properties.

Each recommendation pair was a forced choice, thus the dependent variable was which film the participant would choose and in turn which recommender they would trust (see the screenshot in Fig 4).

Fig 4 Film recommendation.

In addition to the actual choice, we recorded a confidence rating on a Likert scale from 1 to 5 for each choice.

Further, we examined the ratio of the number of times a particular profile was chosen and the number of times a profile was seen (choice ratio).

3.2.5 Results — overall profiles chosen

Figure 5 shows the overall trend of choice ratio, as well as how these profiles were rated in the end in terms of trust.

Trust was elicited through a Likert scale (1—5) in phase 4 in the post-study questionnaire. It is interesting to note the trend shown in the graph in Fig 5. When mapped against profile characteristics, it points to a significant influence of profile similarity and rating overlap.

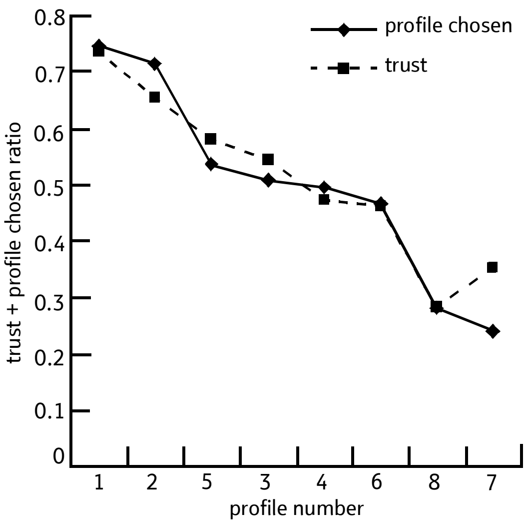

Fig 5 Profile choice ratios/trust rating.

Figure 6 shows the above trend mapped out with the individual profile characteristics.

Overall profile similarity and high rating overlap had an effect on the choice of which recommender to trust.

Participants rated the same profiles in terms of trust in phase 4 of the experiment.

Interestingly, the trend for the trust ratings is the same as the one for the profiles chosen, except for profiles 7 and 8. Here participants trusted the familiar profile more, whereas they chose the unfamiliar one more often.

3.2.6 Individual variable analysis

Since in each condition recommenders differed in one characteristic, while the other characteristics were kept constant, it is worth examining the independent variables individually within each condition type. The results are summarised below and detailed in Fig 7.

profile 1	profile 2	profile 5	profile 3
familiar	unfamiliar	familiar	familiar
similar profile	similar profile	dissimilar profile	similar profile
high rating overlap	high rating overlap	high rating overlap	low rating overlap

profile 4	profile 6	profile 8	profile 7
unfamiliar	unfamiliar	unfamiliar	familiar
similar profile	dissimilar profile	dissimilar profile	dissimilar profile
low rating overlap	high rating overlap	low rating overlap	low rating overlap

Fig 6 Chosen profiles with ratios and characteristics.

familiarity conditions				similarity conditions				rating overlap conditions			
1	profile 1 **55.75%** familiar	vs	profile 2 **44.25%** unfamiliar	5	profile 1 **83.75%** similar	vs	profile 5 **16.25%** dissimilar	9	profile 1 **84.5%** high rating overlap	vs	profile 3 **15.5%** low rating overlap
2	profile 3 **50%** familiar	vs	profile 4 **50%** unfamiliar	6	profile 3 **87.5%** similar	vs	profile 7 **12.5%** dissimilar	10	profile 5 **88.5%** high rating overlap	vs	profile 7 **11.5%** low rating overlap
3	profile 5 **56.75%** familiar	vs	profile 6 **43.25%** unfamiliar	7	profile 2 **84.75%** similar	vs	profile 6 **15.25%** dissimilar	11	profile 2 **85.5%** high rating overlap	vs	profile 4 **14.5%** low rating overlap
4	profile 7 **48.75%** familiar	vs	profile 8 **51.25%** unfamiliar	8	profile 4 **84.75%** similar	vs	profile 8 **15.25%** dissimilar	12	profile 6 **81.5%** high rating overlap	vs	profile 8 **18.5%** low rating overlap

Fig 7 Profiles chosen — individual percentages per condition.

- Familiarity

  In this experiment familiarity did not prove to be a significant influence on participants' choices. Overall only condition 3 turned out to be statistically significant.

- Profile similarity

  Profile similarity had a significant impact on the recommender choice.

- Rating overlap

  Rating overlap also had a significant impact on the recommender choice. Thus participants generally chose films from recommenders with whom they had a high rating overlap. Examining the different visualisations of rating overlap, text and symbol, it is worth noting that there was a higher tendency to choose films from recommenders with high rating overlap in the symbol conditions (87.6 %) rather than in the text conditions (82.5 %).

3.2.7 Choice confidence

For every choice, participants provided a confidence rating on a Likert scale from 1 to 5. We considered a significant deviation of the mean from 3 ($p < 0.05$) as examinable. With the exception of condition 10, all rating means significantly deviated from 3 (see Fig 8).

- Familiarity conditions

  The only condition where participants were confident in their choices was condition 1, where the choice was between a familiar and unfamiliar recommender that was similar and had high rating overlap with the participant.

- Profile similarity conditions

  In those conditions it was interesting to observe that participants were more confident in their choices in conditions 5 and 7 where recommenders had a high rating overlap, than in conditions 6 and

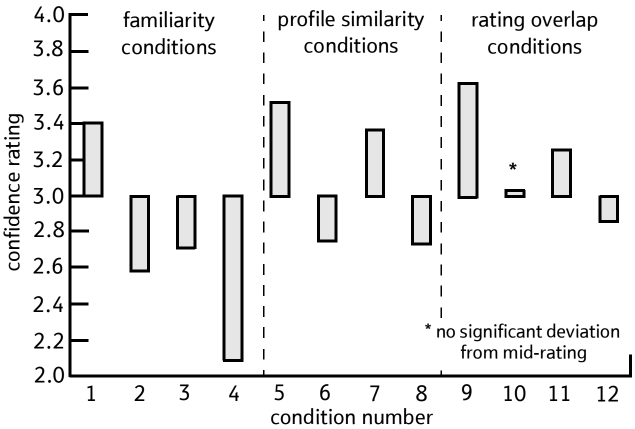

Fig 8 Choice confidence rating means.

8 where recommenders had a low rating overlap with the participant.

- Rating overlap conditions

  Similar to the profile similarity conditions, participants were more confident in their choices in conditions 9 and 11, where the recommenders had a similar profile, compared to conditions 10 and 12, where the recommender profile was dissimilar.

3.2.8 Post-study questionnaire — recommender characteristics

In phase 4 we also asked participants to rate how important the following are (Likert scale, 1—5) when making a movie choice decision in such an on-line recommendation context as well as when they are seeking advice from friends (see Fig 9).

- film title,
- recommender,
- age,
- gender,
- profession,
- hobbies,
- music tastes,
- rating/film taste overlap.

All the ratings significantly deviated from the mid-rating except for the importance of the film title in the on-line context. For the real-world context, the film title was rated as important. In both contexts, participants placed importance on who recommends the film. Age, gender and profession were not rated as particularly important.

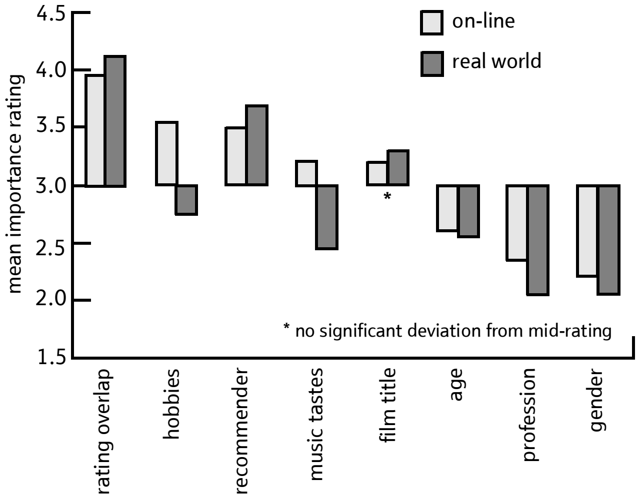

Fig 9 Importance ratings for recommender characteristics (ordered by effect size).

In the on-line context, hobbies were rated as more important than in a real-world context where they were not important. This observation is the same for music tastes in the real world context. Rating/film taste overlap were rated important in both contexts.

3.2.9 Qualitative data

Participants could also include comments about reasons for their decisions in this on-line context, as well as when seeking recommendations in the real world. A total of 151 comments were examined using content analysis [34].

In an on-line context, while there was a universal consensus that rating and film taste overlap was very important, it was revelatory to see that a large number of participants judged profile similarity as an important factor in their decision making. Many of those actually mentioned profile similarity and rating overlap together as an important factor ('... similarity of recommender in terms of hobbies and other films we both enjoyed.')

In a real-world context, participants stated that it mattered how well they knew the recommender ('... how close a recommender is as a friend', '... how close this person is to me, e.g. my boyfriend would influence me more than a distant acquaintance'). Further, whether participants had had past experience with recommendations from a particular person played an important role ('... the recommender's history of good or bad recommendations.')

4. Discussion and conclusions

The results from our studies form the basis for a new perspective on RS design. Since these systems aim to support users in their advice seeking and decision making, their requirements and design should reflect this in supporting users' real-world strategies. The first part of this discussion therefore considers social system design implications, while the second presents an outlook for future research.

4.1 Key conclusions

Our studies have established the following key findings:

- in taste domains, people prefer recommendations from people they know (study 1),
- where the recommenders are not well known, auxiliary information about the recommenders' preferences helps evaluate the appropriateness of recommendations (studies 1 and 2),
- the above can be applied to on-line social recommender environments (study 2).

4.1.1 Recommendations as a social tool

RSs should support the social element of advice seeking through communication and more explicit user matching functions.

When seeking advice from peers in taste domains, people include a variety of social processes of screening potential sources and evaluating the recommendations they receive. In the real world, recommendations are accompanied by a plethora of social and communicative cues, which in an on-line context are missing.

Seeking and receiving a recommendation is a social activity that often involves the discussion of a particular item. Who is consulted for particular advice and why? Why did the recommender like it? Would the recommender want to experience/buy it again? Will the experience change after a while?

RSs should support these social and communicative functions in being more than a mere source of item information. In matching users according to their tastes, collaborative filter RSs are effectively automating the naturally occurring social process of establishing common ground between two strangers.

4.1.2 Social recommender systems

Integrating RSs with social networking functions addresses the shortcomings and capitalises on the advantages of both.

We propose a new breed of social systems which integrate the basic social networking functions of systems such as Facebook or Friendster with those of RSs such as MovieLens. Revealing more about the source of a recommendation, especially the recommender profiles, makes it easier for advice seekers to judge the validity and appropriateness of a recommendation. Users would be able to spot what else that particular recommender has liked in the past and if there are any other elements the advice seeker might have in common.

The main aim of social networking applications is connecting with like-minded people [27]. In current social systems, finding like-minded people still relies on browsing networks of friends, or specifying individual keywords in a database search. In networks of thousands of people this is a very time-consuming, cumbersome and thus ineffective process. Taste-matching, as done through collaborative filter RSs can significantly improve this process.

Profiles in social systems and the information they contain are often perceived as virtual social capital. Users express their individuality through their favourite films, bands, books and TV shows. This already provides a basic binary form of evaluation, which can be qualified further.

User matching could be extended from mere item ratings to further profile information such as interests and preferences. Communication among users, which encourages the questioning of certain recommendations as well as regular networking of like-minded users, would broaden the scope of both classic RSs and social networking applications.

4.1.3 Addressing weaknesses and capitalising on strengths

Social recommender systems address the incentive problems of users, not contributing ratings, and make it easier to find like-minded people.

Integrating social networking and RSs could also address some of the inherent weaknesses of each. RSs suffer from an incentive problem in that, once users have consumed their recommendations, they simply do not contribute any more ratings, which is due to limited utility. If the only purpose of an RS is information retrieval, then the only time people will use it is when they need this type of information, which can be very infrequent.

Social networking applications only provide very cumbersome methods of finding like minded people. Both of these problems can be addressed through a social recommender system on two levels.

- User level

  Matching can be improved through the consideration of classic collaborative filter techniques of matching user with their nearest neighbours. By highlighting not only the items, but also the interests, hobbies and keywords they have in common, the task of finding like-minded users can be made significantly easier.

- Item level

  Recommender algorithms could potentially be extended to consider the keywords that users specify in their profiles. For items such as bands, films or restaurants, these could be directly linked to eCommerce, thus providing an additional basis of common interest and communication.

Overall, a social recommender system addresses the above problems by expanding the utility of classic RSs by becoming a communication platform to stay in touch with friends, as well as improving the networking functionalities of social networking application.

4.1.4 Profile similarity and tagging

Profile characteristics are tags that can be used for establishing profile similarity visualisations and user matching.

In providing rich data about themselves through keywords, whether it is films, bands or TV shows they like, users are effectively providing a descriptive basis about themselves that has already been successfully applied in other domains, i.e.tagging.

Tagging, which is a method of identifying and describing items through keywords, has successfully been applied to Web sites (http://del.icio.us), photos (www.Flickr.com) and music (www.Last.fm). This form of identification does not force a predefined categorisation of items, but rather allows taggers to create their own personal item classification through keywords they find appropriate for an item with the added benefit of indicating preferences.

User profiles in social systems are effectively a form of tagging oneself. Users specify which films and music they like, their interests and hobbies and further keywords that describe them.

Specifically our second study showed that profile similarity visualisations have an impact on participants' choices. We therefore encourage RS algorithm designers to consider how profile data can be used to match users on different levels.

While it is clear that having a common interest in basketball does not necessarily mean that two users will like the same films, depending on how many other interests they might have in common, this information can be useful for judging the appropriateness of a recommendation.

4.2 Future work

4.2.1 Work in progress

Considering that familiarity, as operationalised in study 2 above, did not have any effect, we wanted to examine how real familiarity would affect participants' choices in an on-line RS context. That is why we have designed an experiment similar to the one in study 2 where some of the recommenders would be people the participants actually knew.

Participants are recruited in groups of five under the assumption that they know each other to varying degrees. Familiarity in this experiment is therefore real. To date a pilot has been conducted with 20 participants.

Preliminary results indicate that participants overwhelmingly prefer recommendations from familiar recommenders.

4.2.2 Future research

Further HCI research should examine how communication and networking features can be integrated with RS functionality to deliver a better user experience. Which specific characteristics of a profile are particularly important to consumers and why? How can profile similarity be incorporated into the recommendation strategy?

There is also the question of privacy, i.e. whether users of such systems would be willing to reveal personal information such as interests and preferences. Also, how can the combination of rating overlap and profile similarity information be used to combat malicious attacks on or manipulation of RSs?

On a larger scale, researchers also need to examine how people manage their on-line identities in relation to their real-world identities, especially with regard to how much information they share. Today's teenage generation has adopted social networking systems as a mainstream form of communication. There is, however, one major difference to previous communications tools — information posted can be logged and is often asynchronously available to strangers. Personal communication is therefore not ephemeral any more as private telephone conversations are, but often available and traceable through open discussion boards, blogs, guest-book entries and Internet archives.

With the increasing take-up of social networking applications by the public, social recommender systems are be the next generation of such systems, capable of delivering improved utility and higher user satisfaction.

Precisely because of this increased take-up and its communication potential, these systems need to be researched in order to be able to be designed to deliver to user needs as well as concerns.

Our studies are taking the first steps in that direction in establishing key principles behind people's decision-making and advice-seeking processes and thus making design recommendations based on that. We encourage other researchers to follow this exciting route and address some of the questions above.

References

1 Schwartz B: 'The Paradox of Choice', Harper Collins, (2005).

2 MovieLens — http://movielens.umn.edu/

3 Herlocker J L, Konstan J L, Borchers A and Riedl J: 'Explaining collaborative filtering recommendations', in Proceedings of the ACM 2000 Conference on Computer Supported Cooperative Work, pp 241—250 (2000).

4 Swearingen K and Sinha R: 'Beyond algorithms: An HCI perspective on recommender systems', in Proc of ACM SIGIR 2001 Workshop on Recommender Systems, New Orleans, Lousiana, pp 24—33 (2001).

5 Swearingen K and Sinha R: 'Interaction design for recommender systems', in Proc Interactive Systems (DIS2002), London (June 2002).

6 Sarwar B M, Karypis G, Konstan J A and Riedl J: 'Item-based collaborative filtering recommendation algorithms', in Proc of the Tenth International World Wide Web Conference, pp 285—295 (2001).

7 Breese J S, Heckerman D and Kadie C: 'Empirical Analysis of Predictive Algorithms for Collaborative Filtering', in Proc of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence, pp 43—52 (1998).

8 Canny J F: 'Collaborative Filtering with Privacy via Factor Analysis', in Proc of SIGIR 2002, Tampere, Finland, pp 238—245 (2002).

9 Sinha R and Swearingen K: 'Comparing Recommendations made by Online Systems and Friends', in Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries (2001).

10 Riegelsberger J, Sasse M A and McCarthy J D: 'The Mechanics of Trust: A Framework for Research and Design', in Journal of Human-Computer Studies (in press) (2005).

11 Massa P and Avesani P: 'Trust-aware Collaborative Filtering for Recommender Systems', in Proc. of The International Conference on Cooperative Information Systems (CoopIS), Larnaca, Cyprus (October 2004).

12 O'Donovan J and Smyth B: 'Trust in Recommender Systems', in Proc of The International Conference on Intelligent User Interfaces, San Diego, California, pp 167—174 (January 2005).

13 Briggs P, De Angelo A, Burford B and Lynch P: 'Trust in Online Advice', in Social Science Computer Review, 20, No 3, pp 321—332 (2002).

14 Bonhard P and Sasse M A: ' 'I thought it was terrible and everyone else loved it' — A New Perspective for Effective Recommender System Design', Springer Verlag, in Proc of the 19th British HCI Group Annual Conference, Napier University, Edinburgh, UK, pp 251—265 (September 2005).

15 Eysenck M W: 'Language and Thought', in: 'Psychology — A Student's Handbook', Hove, Psychology Press Ltd, (2000).

16 Gross R: 'Problem-solving, Decision-making and Artificial Intelligence', in: 'Psychology — The Science of Mind and Behaviour', London, Hodder & Stoughton (2001).

17 Harvey N and Fischer I: 'Taking Advice: Accepting Help, Improving Judgement, and Sharing Responsibility', in Organizational Behavior and Human Decision Processes, 70, No 2, pp 117—133 (1997).

18 Yaniv I: 'Receiving other peoples' advice: Influence and Benefit', Organizational Behavior and Human Decision Processes, 93, No 1, pp 1—13 (2004).

19 Yaniv I: 'The Benefit of Additional Opinions', American Psychological Society, 13, No 2, pp 75—78 (2004).

20 Yaniv I and Kleinberger E: 'Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation', Organizational Behavior and Human Decision Processes, 83, No 2, pp 260—281 (2000).

21 Perugini S, Goncalves M A and Fox E A: 'A connection centric survey of recommender system research', Journal of Intelligent Information Systems, 23, No 1 (2004).

22 Lueg C: 'Social Filtering and Social Reality', ERCIM, in Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering, Budapest, Hungary, pp 73—77 (November 1997).

23 Friendster — http://www.friendster.com/

24 Facebook — http://www.facebook.com/

25 hi5 — http://www.hi5.com/

26 mySpace — http://www.myspace.com/

27 Donath J and boyd d: 'Public displays of connection', BT Technology Journal, 22, No 4, pp 71—82 (October 2004).

28 boyd d and Heer J: 'Profiles as Conversation: Networked Identity Performance on Friendster', in Proceedings of the Hawaii International Conference on System Sciences (HICSS-39), Kauai, HI: IEEE Computer Society, (January 2006).

29 Berscheid E and Reis H T: 'Attraction and close relationships', in Gilbert D T, Fiske S T and Linzey G (Eds): 'The Handbook of Social Psychology', 2, 4th, 22, Oxford University Press, pp 193—281 (1998).

30 Strauss A and Corbin J: 'Basics of Qualitative Research — Techniques and Procedures for Developing Grounded Theory', London, SAGE Publications (1998).

31 Perugini S, Goncalves M A and Fox E A: 'Recommender Systems Research: A Connection-Centric Survey', in Journal of Intelligent Information Systems, 23, No 2, pp 107—143 (2004).

32 Bonhard P, Harries C, McCarthy J D and Sasse M A: 'Accounting for Taste: Using Profile Similarity to Improve Recommender Systems', In Proceedings of CHI 2006, Montreal, Canada, pp 22—27 (April 2006).

33 Gigerenzer G: 'Fast and Frugal Heuristics: The Tools of Bounded Rationality', in Harvey N and Koehler D (Eds): 'Blackwell Handbook of Judgment and Decision Making', Oxford, Blackwell, 4, pp 62—88 (2004).

34 Flick U: 'An Introduction to Qualitative Research', London, Sage Publications (1998).

Philip Bonhard is a PhD candidate in the Human Centred Systems Group at University College London. He specialises in researching advice-seeking and decision-making behaviour in on-line contexts, notably with respect to recommender systems and social networking applications.

Previously he studied European politics at King's College London and the London School of Economics and subsequently Computer Science at UCL.

Angela Sasse is a Professor for Human-Centred Technology in the Department of Computer Science at UCL.

She has been researching design and usability issues of novel communication technologies since joining UCL as a lecturer in 1990. She read psychology in Germany before obtaining an MSc in Occupational Psychology from the University of Sheffield, and a PhD in Computer Science from the University of Birmingham.