UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
You're special, but it doesn't matter if you're a greenhorn: Social recommender strategies for mere mortals

Permalink
https://escholarship.org/uc/item/7gm7k27n

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors
Analytis, Pantelis P
Barkoczi, Daniel
Herzog, Stefan M

Publication Date
2015

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

You're special, but it doesn't matter if you're a greenhorn: Social recommender strategies for mere mortals

Pantelis P. Analytis (analytis@mpib-berlin.mpg.de) **Daniel Barkoczi (barkoczi@mpib-berlin.mpg.de)**
Center for Adaptive Behavior and Cognition (ABC), Max Planck Institute for Human Development
Lentzeallee 94, 14195, Berlin, Germany

Stefan M. Herzog (herzog@mpib-berlin.mpg.de)
Center for Adaptive Rationality (ARC), Max Planck Institute for Human Development
Lentzeallee 94, 14195, Berlin, Germany

Abstract

From choosing a book to picking a restaurant, most choices people encounter are about "matters of taste" and thus no universal, objective criterion about the options' quality exists. Tapping into the knowledge of individuals with similar tastes who have already experienced and evaluated options—as harnessed by recommender system algorithms—helps people select options that they will enjoy. Although recommender systems are available in some domains, for most everyday decisions there is neither an algorithm nor "big data" at hand. We mapped recommender system algorithms to models of human judgment and decision making about "matters of fact" and then recast the latter as social recommender strategies for "matters of taste". This allowed us to investigate how people can leverage the experiences of other individuals to make better decisions when no machine recommender systems are available. Using computer simulations on a widely used data set from the recommender systems literature, we show that *experienced* individuals can benefit from relying on only the opinions of seemingly similar people. *Inexperienced* individuals, in contrast, are often well-advised to pick the mainstream option (i.e., the one with the highest average evaluation) even if there are interindividual differences in taste; this is because reliable estimation of similarity requires considerable experience.

Keywords: Social learning; wisdom of crowds; expert crowd, recommender systems; learning.

Introduction

Where should I go for my next vacation? Which car should I buy? Most choices people encounter are about "matters of taste" and thus no universal, objective criterion about the options' exists. How can people increase their chances of selecting options that they will enjoy?

One promising approach is to tap into the knowledge of other individuals who have already experienced and evaluated options. The recommender systems community has leveraged this source of knowledge to develop *collaborative filtering methods*, which estimate the subjective quality of options that people have not yet experienced (Resnick & Varian, 1997; Adomavicius & Tuzhilin, 2005). One key insight is that building recommendations based only on the evaluations of individuals *similar* to the target individual often improves the quality of the recommendations (e.g., Herlocker, Konstan, Borchers, & Riedl, 1999)—where similarity between two people is typically defined as the correlation in their evaluations across options they have both evaluated.

Although the consumer industry enables people to benefit from recommender systems in some domains (e.g., choosing a movie on Netflix), for many everyday decisions there is neither an algorithm nor "big data" at hand. How can individuals leverage the experience of other people when they have no access to big data but access only to a relatively small community of other people with whom they share some prior experience about the available options?

In this paper we make three contributions. First, we have undertaken an exercise in *theory integration* by mapping the striking conceptual similarities between seminal recommender system algorithms and both (i) models of judgment and categorization and (ii) models of social learning and social decision making (from psychology, cognitive science, judgment and decision making, anthropology, and biology). Second, we have recast the latter two classes of models as *social recommender strategies*. Finally, based on this mapping, we have investigated how ordinary people can leverage the experience of other people to make better decisions about matters of taste. To this end we studied the inevitable trade-off between (i) harnessing the apparent (dis)similarity between people's tastes—to discriminate between more and less relevant advisers—and (ii) estimating those similarities accurately enough. We have investigated how this trade-off evolves with the amount of experience a decision maker has (i.e., the number of options previously evaluated).

Outside of the recommender systems literature, social recommender strategies remain an under-explored topic. Research on advice taking, social learning, and judgment aggregation in psychology, cognitive science, judgment and decision making, anthropology, and biology has focused almost exclusively on "matters of fact" where there is an objective criterion to be inferred ("wisdom of crowds"; e.g., Larrick, Mannes, & Soll, 2012). To the best of our knowledge, there are only a handful of studies on social recommender strategies (Van Swol, 2011; Yaniv, Choshen-Hillel, & Milyavsky, 2011; Müller-Trede, Choshen-Hillel, Barneron, & Yaniv, 2015). They show that people rely on the similarity between themselves and their advisers when making decisions about matters of taste and that this is a good strategy.

Mapping recommendation systems algorithms to *informational* and *social* cue-based strategies

Table 1 displays several social recommender strategies that predict one's own future evaluations based on the past evaluations provided by other people.

	Social recommender strategy		Parallels in the literature		
Strategy	Verbal description	Formal definition	Informational cues	Social cues	Recommender systems
Follow your Doppelgänger (cf. Yaniv et al., 2011)	Find individual s with the most similar taste and adopt that individual's evaluations as your own estimates.	$\hat{u}_i = u_s$	*Take the best* (Gigerenzer & Goldstein, 1996), *single attribute*, (Hogarth & Karelaia, 2005)	*Imitate the best* (Richerson & Boyd, 2008)	*Nearest neighbors* ($k = s = 1$)
Follow the whole crowd	Average evaluations of all N other individuals (i.e., go with the mainstream).	$\hat{u}_i = 1/N \times \sum_{j=1}^{N} u_j$	*Equal/unit weights* (Dawes, 1979)	*Averaging* (Einhorn, Hogarth, & Klempner, 1977)	*Nearest neighbors* ($k = N$), often used as a benchmark (e.g., Shardanand & Maes, 1995)
Follow your clique	Average evaluations of the k most similar individuals.	$\hat{u}_i = 1/k \times \sum_{j=1}^{k} u_j$	–	*Select crowd* (Mannes et al., 2014), *expert crowd* (Goldstein et al., 2014)	*Nearest neighbors* ($1 < k < N$; Shardanand & Maes, 1995)
Follow your similar crowd	Average evaluations of all k individuals whose taste is correlated with yours above a similarity threshold t.	$\hat{u}_i = 1/k \times \sum_{j=1}^{k} u_j$	–	–	Common implementation of *nearest neighbors* (Desrosiers & Karypis, 2011).
Follow the similarity-weighted crowd (cf. Müller-Trede et al., 2015)	Weight evaluations of all N individuals according to their similarity to your taste.	$\hat{u}_i = \frac{1}{\sum_{j=1}^{N} w_j} \sum_{j=1}^{N} w_j \times u_j$	*Weighted average* (Hammond, Hursch, & Todd, 1964; Dana & Dawes, 2004)	*Weighted crowd* (Davis-Stober et al., 2014)	*Weighted neighbors* (Resnick, Iacovou, Suchak, Bergstrom, & Riedl, 1994)
Consider similar options	Find k most similar options (i.e., with similar evaluation profiles across people) and weight your own evaluations for them according to their similarity.	$\hat{u}_i = \frac{1}{\sum_{l=1}^{k} w_l} \sum_{l=1}^{k} w_l \times u_l$	*Exemplar models* (Kruschke, 1992; Juslin & Persson, 2002)	–	*Item-based algorithms* (Sarwar, Karypis, Konstan, & Riedl, 2001)
Follow random other (cf. Gilbert et al., 2009)	Select an individual r at random and adopt that individual's evaluations as your own estimates.	$\hat{u}_i = u_r$	*Minimalist* (Gigerenzer & Goldstein, 1996)	*Random copying* (Cavalli-Sforza, 1981)	Occasionally used as benchmark strategy

Table 1: Social recommender strategies conceptually similar to strategies using informational cues or social cues (i.e., people's opinions as cues). Due to limited space we report only representative references. All strategies first estimate the expected utility $\hat{u}_i$ (i.e., enjoyment) of each option i and then select the option with the highest estimated utility; when several options have the same estimated utility, one of the tied options is chosen at random. Strategies incorporating similarity information are typeset in *italics* and those averaging across several individuals' evaluations are typeset in **bold**. All strategies are person based, except consider similar options, marked in *blue*.

We identified conceptual similarities between the proposed social recommender strategies (inspired by seminal algorithms from recommender systems research) and heuristics and strategies for predicting matters of fact, where people have access to either *informational cues* potentially related to an objective criterion (e.g., number of movie theaters in a city to predict its population size) or *social cues* (i.e., the opinions of other people concerning the same objective criterion).

This mapping emphasizes the close correspondence between recommendation algorithms on the one hand and informational and social cue-based strategies on the other. The social recommender strategies can be placed in a continuum, on the boundaries of which strategies rely either only on similarity information or only on aggregation of opinions (see Table 1). As we move away from the boundaries the strategies rely increasingly on both these two fundamental principles. Below we illustrate some of the strategies using a fictional data set that has the same structure as the large-scale data sets used in recommender systems research and in our own simulation study below.

Example: Deciding which movie to watch based on other people's past experiences

Amit likes superhero movies and wants to watch *Batman* or *Fantastic Four*. His friends have already seen both movies. Furthermore, he and his friends have all watched and evaluated several other movies (see Table 2). In addition to any other contextual information (e.g., director, cast, movie length), Amit can use his friends' evaluations to inform his movie choice.

Movie	John	Bob	Linda	Mary	Lou	*Avg.*	**Amit**
Superman	3	4	2.5	4.5	3	*3.8*	**2.5**
Spiderman	4	4.5	3	2	3.5	*3.4*	**3**
Batman	5	5	2	1	3	*3.2*	**?**
Fantastic Four	2	3	2.5	3	2.5	*2.6*	**?**
X-Men	1	1.5	2	1.5	3	*1.8*	**2**

Table 2: A typical recommender system problem. The movies are rated on a scale of 1 to 5 (higher values indicate more positive ratings). *Avg. = average*.

From Amit's perspective, his own future evaluations are the criterion values he seeks to maximize and the evaluations of his friends are informational cues he can use to predict his own future evaluations. Based on his past evaluations of the other movies, Amit thinks that he and Linda have similar taste. If Linda truly were his "taste Doppelgänger" he could simply copy her evaluations and arrive at very accurate estimates of his own future enjoyment (*Follow your Doppelgänger*). However, it is unclear to what extent this seeming similarity—based on only a small set of joint past experiences—would generalize well to future cases. Amit may thus prefer to take the evaluations of others into account, as well. For example, he could assign equal weights to all individuals and simply use the average evaluation (i.e., the "mainstream" option; *Follow the whole crowd*). Yet by doing so he would also incorporate evaluations from individuals with possibly very different—or even antithetical—tastes. Alternatively, he could search for a movie that everybody rated similarly to the target movie and then use his own evaluation for that similar movie as a proxy (*Consider similar options*; e.g., *Spiderman* is similar to *Batman*).

Simulation study

We investigated the performance of the proposed social recommender strategies (see Table 1) by simulating their predictions for a large-scale, real-world data set. We varied the experience of the simulated decision makers (i.e., the number of options previously experienced in that domain; that is, the number of rows in Table 2). As experience increased, the strategies relying on similarity could thus base their similarity estimates on more data.[1]

The social network from which a person could leverage vicarious experience would likely be much smaller than the thousands of people available in typical recommender system data sets. The cognitive limit of the number of stable relationships that people can maintain is estimated to be around 250 (Dunbar, 2010). We therefore opted to simulate small "communities" of 250 members each to mirror this real-world feature (as opposed to letting decision makers have access to all other individuals in the population).

Method

Dataset We used the funniness ratings of 100 jokes collected in the *Jester* data set. Jester[2] was created by an online recommender system that allows Internet users to read and rate jokes. Users evaluated jokes on a scale ranging from *not funny* (–10) to *funny* (+10). At the beginning of the recommendation process, a set of 10 jokes was presented to the user. Thereafter, Jester recommended jokes and continued to collect ratings for each of them. The data set contains 4.1 million evaluations of 100 jokes by 73,421 participants. In contrast to other data sets studied by the recommender system community, here a large number of participants evaluated all options. Since its publication, the Jester data set has been used extensively to study collaborative filtering algorithms.

Simulation procedure For simplicity we worked only with participants who evaluated all jokes (reducing the number of participants from 73,421 to 14,116). We randomly selected 14,000 participants in order to partition them into evenly sized communities of 250 members each. In line with previous work in the recommender system literature, we used the Pearson correlation coefficient as a measure of similarity (Herlocker, Konstan, Terveen, & Riedl, 2004).[3]

In each simulation run, we followed the following steps:

[1] A similar challenge is faced by recommender system algorithms when recommending options to new users about whom they know nothing or only very little. This challenge is commonly referred to as the *user cold start problem* (Ekstrand, Riedl, & Konstan, 2011).

[2] http://eigentaste.berkeley.edu

[3] The Pearson similarity coefficient between two individuals or two items i and j is defined as $w(i,j) = \frac{\sum_{n=1}^{k}(u_{in}-\bar{u_i})(u_{jn}-\bar{u_j})}{\sum_{n=1}^{k}\sqrt{(u_{in}-\bar{u_i})^2(u_{jn}-\bar{u_j})^2}}$

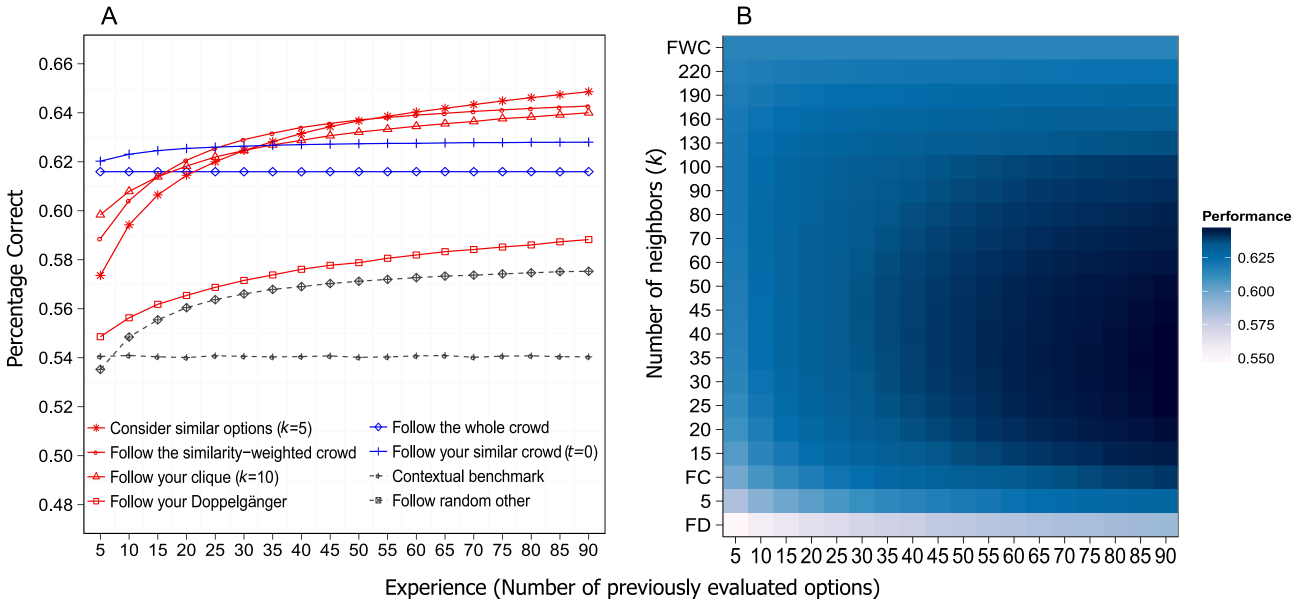

Figure 1: **Panel A**: The performance of strategies as a decision maker's experience with the domain of jokes (i.e., number of jokes previously experienced and evaluated) increases; the strategies are grouped by color to those that rely primarily on aggregation (blue), those that rely heavily on similarity information (red) and benchmark strategies (black) (see also Table 1). **Panel B**: Performance of the *Follow your clique* strategy as a function of the experience with the domain (i.e., number of options experienced; x axis) and the size of the clique (i.e., number of most similar people consulted; y axis). Note that *Follow your Doppelgänger* and *Follow the whole crowd* are special cases of this strategy when the number of similar people consulted equals 1 and N, respectively. **FD**: *Follow your Doppelgänger*. **FC**: *Follow your clique*. **FWC**: *Follow the whole crowd*.

First, we randomly generated 56 communities with 250 members each (14,000/250). Second, we randomly divided the jokes into a training (x jokes) and a test (10 jokes) set; this assignment was the same for all individuals within all communities. The strategies were then fitted on the training set. Individuals could access only advisers within their own community. Third, for each individual (within all communities) we generated all 45 possible pair comparisons within the test set $[10 \times (10-1) \div 2]$ and examined the performance of the strategies in predicting which of the two jokes in a pair had a higher evaluation for that individual, resulting in 45 pair comparisons per individual, 11,250 per community (45×250), and 630,000 in total ($11,250 \times 56$). For each strategy we recorded the proportion of correct predictions. This procedure was repeated 100 times and results were averaged. We investigated how the performance of the strategies changed as a function of experience by repeating this procedure for different numbers (x) of jokes experienced in the training set (varying from 5 to 90 in increments of 5).

Results

How did the strategies perform? Figure 1 shows the performance of each strategy as a function of the number of options evaluated. For the highest level of experience the strategy based on item similarity (*Consider similar options*) performed best (predicting 65% of the pair comparisons correctly). This was followed by the strategies that relied on both similarity information and aggregation: *Follow your clique* and *Follow the similarity-weighted crowd* predicted approximately 64% of the cases correctly and *Follow your similar crowd*—relying on similarity information more crudely—performed slightly worse at 63%. Strategies relying solely on either user similarity (*Follow your Doppelgänger*) or aggregation (*Follow the whole crowd*) performed worse than the other strategies, reaching 59% and 62%, respectively.

The usefulness of less similar advisers These results provide a rationale for why people rely on similar advisers (Yaniv et al., 2011; Müller-Trede et al., 2015). However, relying purely on similarity (*Follow your Doppelgänger*) does not perform that well because of the difficulty of reliably estimating similarity in light of sampling error. Mirroring results from research on the wisdom of crowds (Goldstein et al., 2014; Mannes et al., 2014), taking into account additional—although less similar—advisers and averaging their recommendations markedly improves performance.

Experience within a domain Strategies that rely heavily on similarity information have steep learning curves. For small amounts of experience, *Follow your similar crowd* is the best performing strategy. *Consider similar options*, *Follow your clique*, and *Follow the similarity-weighted crowd* start to outperform *Follow the whole crowd* once approximately 15 options have been added to the training set and *Follow your similar crowd* after approximately 25 options.

Thus, decision makers who have not yet experienced many options are well-advised simply to aggregate the evaluations of individuals who seem to have at least minimally similar (i.e., positively correlated) tastes (*Follow your similar crowd*) or even to unconditionally aggregate the evaluations of all individuals (*Follow the whole crowd*). Although the opinions of truly similar individuals are more informative than those of truly dissimilar individuals, this discrimination is only beneficial to the extent that it is accurate enough. For small training samples, estimates of similarity are apparently often not accurate enough to be of any use.

How large should your clique be? When *following your clique*, the size of k (i.e., the number of neighbors whose evaluations are averaged) is a *hyperparameter* that needs to be chosen beforehand (in Figure 1A we fixed $k = 10$). Figure 1B shows how performance changes as a function of k and experience (i.e., the number of options experienced). With little experience it is better to rely on large cliques (ca. 100), whereas for more extensive experience, performance peaks at moderately sized cliques (ca. 30).

The potential of one-reason decision making Following a *random other* person correctly predicted 54% of the comparisons, which indicates a very modest, minimally shared sense of humor in the population (i.e., slightly better than chance). We also tested another benchmark strategy that simply used the length of a joke as a cue to predict its evaluation (i.e., some people may prefer long, story-like jokes while others may prefer short and witty linguistic puns). This one-cue strategy predicted 57% of cases correctly—it was almost as accurate as the *Follow your Doppelgänger* strategy (59%), which also relied on one cue, yet a social one.[4]

General discussion

Mapping out the striking conceptual similarities between seminal recommender system algorithms, on the one hand, and extant models of judgment and decision making (based on informational or social cues), on the other, allowed us to recast the latter models as social recommender strategies (see Table 1). This theory integration allowed us to analyze the performance of social recommender strategies for *mere mortals* who have access to only a small pool of potential advisers, rather than the "big data" available to recommender systems.

Two results stand out. First, the successful strategies all have one thing in common: They aggregate evaluations across several people (or items). Second, the amount of experience within a domain turns out to be a crucial determinant of the success of strategies using similarity information. Whereas *experienced* people can benefit from relying on only the opinions of seemingly similar people, *inexperienced* people are often well-advised to aggregate the evaluations of a large set of people (picking the option with the highest average evaluation either across all people or across at least minimally similar people) even if there are interindividual differences in taste, because reliable estimation of similarity requires considerable experience.

[4]This result conflicts with a relevant finding from a speed dating experiment (Gilbert et al., 2009), where the experience of a random other person (from the same population) predicted the actual dating enjoyment better than the same participants' predictions (based on an extensive set of informational cues available before the speed date started, namely, among other things, a picture and information about age, height, favorite movie, sport, book, song, and food).

Experience and the bias–variance trade-off

With increasing experience with the domain, the performance of all top-notch strategies increased—except for the wisdom of crowds strategy (*Follow the whole crowd*), which unconditionally averages across all people and is thus—by design—unaffected by the increasing accuracy of the similarity estimates. Such an averaging strategy assumes that everybody has the same taste and performs well to the extent that the tastes in the population are indeed homogeneous. From a bias–variance trade-off perspective (e.g., Gigerenzer & Brighton, 2009; Geurts, 2010), this strategy suffers from potentially high bias to the extent that its homogeneity assumption is wrong, but exhibits zero variance in its prediction error because it does not estimate any free parameters.[5]

In contrast, the strategies relying on similarity have a comparatively low bias because they can adapt to the homogeneity or heterogeneity of tastes in the population. However, they potentially suffer from variance because their predictions depend on the similarity estimates—to differing degrees—and thus they lie on a bias–variance continuum. At one extreme, a strategy of adopting the evaluations of only the seemingly most similar person has the potential to profit from the vicarious experiences of one's taste Doppelgänger but is most reliant on an accurate estimation of similarity. At the other extreme, a strategy of relying on a large crowd of at least minimally similar people (i.e., with at least positively correlated tastes) is more biased but also more robust because it depends on only roughly discriminating between similar and dissimilar advisers (see also Goldstein et al., 2014; Mannes et al., 2014).

Theory integration: Reconnecting the cognitive sciences with recommender systems research

New statistical tools haven often served as an inspiration for the development of new psychological theories (Gigerenzer, 1991). In the case of recommender systems, however, the insights developed within the last two decades have not been much incorporated into cognitive science[6]—despite recommender systems being widely available and relevant for everyday decision making and seminal recommender systems being inspired by the work of cognitive scientists (Rich, 1979). We hope that the current paper initiates a cross-fertilization between the two until now largely unconnected research streams.

Context-based and hybrid recommender strategies

In everyday life, people and machines have access to information beyond their own and other people's past experiences: informational cues describing options and advisers (e.g., a movie's genre and a person's clothing style, respectively).

[5]Also from a Bayesian perspective, it is prudent to go with the crowd: An inexperienced decision maker—by statistical necessity—is a priori more likely than not to have "mainstream taste" unless there is diagnostic private information to the contrary (Herzog & Hertwig, 2013, p. 210).

[6]In a similar vein, Analytis et al. (2014) pointed out the overlooked analogy between ranking models from machine learning and human search behavior.

The use of such contextual information has been examined in multiple-cue judgment and categorization learning in cognitive science, in context-based recommender systems, and, more generally, in supervised learning in machine learning.

People might thus improve their predictions about matters of taste by using informational cues (i) to take advantage of the predictive information in the options' features themselves (e.g., it's a superhero movie) or (ii) to improve their assessment of advisers' similarities (e.g., by looking at their clothing style). Such context-based—or even hybrid—approaches might further improve people's ability to make good decisions for themselves by taking advantage of different sources of information and different approaches (see also Herzog & Hertwig, 2009; Herzog & von Helversen, 2013; Herzog & Hertwig, 2014).

Acknowledgments

We would like to thank the members of the ABC group for their constructive comments during internal presentations and Anita Todd for editing the manuscript.

References

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, *17*(6), 734–749.

Analytis, P. P., Kothiyal, A., & Katsikopoulos, K. (2014). Multi-attribute utility models as cognitive search engines. *Judgment and Decision Making*, *9*(5), 403–419.

Cavalli-Sforza, L. L. (1981). *Cultural transmission and evolution: A quantitative approach* (No. 16). Princeton University Press.

Dana, J., & Dawes, R. M. (2004). The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics*, *29*(3), 317–331.

Davis-Stober, C. P., Budescu, D. V., Dana, J., & Broomell, S. B. (2014). When is a crowd wise? *Decision*, *1*(2), 79–101.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*(7), 571–582.

Desrosiers, C., & Karypis, G. (2011). A comprehensive survey of neighborhood-based recommendation methods. In *Recommender systems handbook* (pp. 107–144). Springer.

Dunbar, R. (2010). *How many friends does one person need?: Dunbar's number and other evolutionary quirks*. Faber & Faber.

Einhorn, H. J., Hogarth, R. M., & Klempner, E. (1977). Quality of group judgment. *Psychological Bulletin*, *84*(1), 158–172.

Ekstrand, M. D., Riedl, J. T., & Konstan, J. A. (2011). Collaborative filtering recommender systems. *Foundations and Trends in Human-Computer Interaction*, *4*(2), 81–173.

Geurts, P. (2010). Bias vs variance decomposition for regression and classification. *Data Mining and Knowledge Discovery Handbook*, 733–746.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, *98*(2), 254–267.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, *1*(1), 107–143.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*(4), 650–669.

Gilbert, D. T., Killingsworth, M. A., Eyre, R. N., & Wilson, T. D. (2009). The surprising power of neighborly advice. *Science*, *323*(5921), 1617–1619.

Goldstein, D. G., McAfee, R. P., & Suri, S. (2014). The wisdom of smaller, smarter crowds. In *Proceedings of the fifteenth ACM conference on economics and computation* (pp. 471–488).

Hammond, K. R., Hursch, C. J., & Todd, F. J. (1964). Analyzing the components of clinical inference. *Psychological Review*, *71*(6), 438–456.

Herlocker, J. L., Konstan, J. A., Borchers, A., & Riedl, J. (1999). An algorithmic framework for performing collaborative filtering. In *Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval* (pp. 230–237).

Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, *22*(1), 5–53.

Herzog, S. M., & Hertwig, R. (2009). The wisdom of many in one mind: Improving individual judgments with dialectical bootstrapping. *Psychological Science*, *20*(2), 231–237.

Herzog, S. M., & Hertwig, R. (2013). The ecological validity of fluency. In C. Unkelbach & R. Greifeneder (Eds.), *The experience of thinking: How the fluency of mental processes influences cognition and behavior* (pp. 190–219). London: Psychology Press.

Herzog, S. M., & Hertwig, R. (2014). Harnessing the wisdom of the inner crowd. *Trends in Cognitive Sciences*, *18*(10), 504–506.

Herzog, S. M., & von Helversen, B. (2013). Blending and choosing within one mind: Should judgments be based on exemplars, rules, or both? In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Cooperative minds: Social interaction and group dynamics. Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 2536–2541). Austin, TX: Cognitive Science Society.

Hogarth, R. M., & Karelaia, N. (2005). Ignoring information in binary choice with continuous variables: When is less "more"? *Journal of Mathematical Psychology*, *49*(2), 115–124.

Juslin, P., & Persson, M. (2002). PROBabilities from EXemplars (PROBEX): A "lazy" algorithm for probabilistic inference from generic knowledge. *Cognitive Science*, *26*(5), 563–607.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22–44.

Larrick, R. P., Mannes, A. E., & Soll, J. B. (2012). The social psychology of the wisdom of crowds. In J. I. Krueger (Ed.), *Frontiers in social psychology: Social judgment and decision making* (pp. 227–242). New York, NY: Psychology Press.

Mannes, A. E., Soll, J. B., & Larrick, R. P. (2014). The wisdom of select crowds. *Journal of Personality and Social Psychology*, *107*(2), 276–299.

Müller-Trede, J., Choshen-Hillel, S., Barneron, M., & Yaniv, I. (2015). The wisdom of crowds for matters of taste. *Manuscript submitted for publication*.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). Grouplens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 acm conference on computer supported cooperative work* (pp. 175–186).

Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, *40*(3), 56–58.

Rich, E. (1979). User modeling via stereotypes. *Cognitive Science*, *3*(4), 329–354.

Richerson, P. J., & Boyd, R. (2008). *Not by genes alone: How culture transformed human evolution*. University of Chicago Press.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on world wide web* (pp. 285–295).

Shardanand, U., & Maes, P. (1995). Social information filtering: Algorithms for automating "word of mouth". In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 210–217).

Van Swol, L. M. (2011). Forecasting another's enjoyment versus giving the right answer: Trust, shared values, task effects, and confidence in improving the acceptance of advice. *International Journal of Forecasting*, *27*(1), 103–120.

Yaniv, I., Choshen-Hillel, S., & Milyavsky, M. (2011). Receiving advice on matters of taste: Similarity, majority influence, and taste discrimination. *Organizational Behavior and Human Decision Processes*, *115*(1), 111–120.