Trust-Related Effects of Expertise and Similarity Cues in Human-Generated Recommendations

Johannes Kunkel, Tim Donkers, Catalin-Mihai Barbu, Jürgen Ziegler
University of Duisburg-Essen
Duisburg, Germany
{johannes.kunkel, tim.donkers, catalin.barbu, juergen.ziegler}@uni-due.de

ABSTRACT
A user's trust in recommendations plays a central role in the acceptance or rejection of a recommendation. One factor that influences trust is the source of the recommendations. In this paper we describe an empirical study that investigates the trust-related influence of social presence arising in two scenarios: human-generated recommendations and automated recommending. We further compare visual cues indicating the expertise of a human recommendation source and its similarity with the target user, and evaluate their influence on trust. Our analysis indicates that even subtle visual cues can signal expertise and similarity effectively, thus influencing a user's trust in recommendations. These findings suggest that automated recommender systems could benefit from the inclusion of social components—especially when conveying characteristics of the recommendation source. Thus, more informative and persuasive recommendation interfaces may be designed using such a mixed approach.

ACM Classification Keywords
H.1.2 Models and Principles: User/Machine Systems—*Human factors*; H.3.3. Information Storage and Retrieval: Information Search and Retrieval—*Information filtering*; H.5.2 Information Interfaces and Presentation (e.g. HCI): User Interfaces—*Evaluation/methodology, Graphical user interfaces (GUI), Theory and methods*

Author Keywords
Recommender Systems; Trust; Personal Recommendation Sources; Structural Equation Modeling

INTRODUCTION
Trust in recommendations is a key aspect in the user's intention to follow them [32, 2]. One of the factors that influences the perceived trustworthiness of recommendations is the source, i.e. the person or system providing them [31, 6, 34]. In online situations, recommendations are typically generated by automated *Recommender Systems* (RS), which are a commonly-used instrument to tackle the information overload users encounter on the Internet. However, platforms where humans provide advice and recommendations to users also exist. There is an increasing number of systems where both automated recommendations and human-generated recommendations, e.g. in the form of online reviews, are available. This raises interesting research questions, such as which recommendation sources are considered more trustworthy by users and what are the factors that influence users' trust.

It has been shown that the bare social presence, which emerges in situations where humans provide recommendations, can already have a positive impact on trust [6]. Trust further depends on perceived traits of the recommendation source, such as expertise and similarity towards the receiver of recommendations [14]. In addition, relying on users with strong social ties to the active user when generating recommendations has shown great potential in fostering system-based trust [30, 33]. However, human recommendation givers on online platforms are typically not known to the user. Thus, an important question in cases where recommendation sources are anonymous is which information about the recommendation source can affect—either positively or negatively—the user's trust.

In this paper, we present our ongoing work regarding trust in recommendation sources with socially-weak ties. In particular, we seek to answer the following research questions:

RQ1: Is the depiction of an anonymous user avatar for the recommendation source sufficient to cause a perception of social presence?
RQ2: Can the expertise and similarity of a recommendation source be conveyed successfully through subtle visual cues?
RQ3: Do the perceived expertise and similarity of an anonymous human recommendation source influence trust in recommendations?

We present the results of an empirical study in which we compared the effects of a recommendation interface, which claimed to show users personalized recommendations provided either by a human or an automated RS, on the perception of social presence. Furthermore, we altered visual cues slightly by showing (1) an iconic face when a human was the recommendation source, and (2) arrows pointing up/down to indicate a high/low expertise or similarity. Our results show that, although the visual elements we used were rather subtle, they were mostly able to influence users' perceptions regar-

HUMANIZE '18, March 11, 2018, Tokyo, Japan

ding the recommendation source's social presence, expertise, and similarity.

To summarize, in this paper we contribute to the research field of trust in recommendation sources by:

- presenting results of a mid-scale survey ($n = 88$) that compares the user's perception of other users and automated RS as recommendation source;
- introducing a structural equation model incorporating the user's perceived similarity and expertise of the recommendation source and their influences on trust; and
- discussing implications for future interface design to foster trust in RS.

The following sections are organized as follows: First, a review of related research is given. Subsequently, we introduce our method for comparing the nature and properties of recommendation sources. Our results are presented and discussed thereafter. Finally, we conclude by summarizing the implications derived from our results, the limitations, as well as further research directions.

RECOMMENDATION SOURCES

Besides aspects such as the presentation of recommendations and their accuracy in meeting the user's preferences, the source of a recommendation and how the user perceives it can also influence the acceptance of recommendations [30, 29, 31, 33]. A general way to differentiate between recommendation sources is to describe them as either *personal* or *impersonal* [29, 31]. Personal sources are directly bound to human beings as provider of recommendations. An example for recommendations by personal sources are *word of mouth* recommendations [8, 14, 19]. Impersonal sources, on the other hand, are automated systems or sources that are not explicitly communicated to the user.

Personal and impersonal sources can provide personalized or non-personalized recommendations. Thus, four recommendation categories can be differentiated [29]: a personal source providing personalized recommendations (e.g. recommendations by friends), a personal source providing non-personalized recommendations (e.g. user product reviews), an impersonal source providing personalized recommendations (e.g. typical RS), and an impersonal source providing non-personalized recommendations (e.g. a public product advertisement).

Recommendations from a personal source can be further subdivided based on their *tie strength*. The tie strength between social peers describes the strength of their relation and the degree to which they know each other [20, 19]. Prior research in personal recommendation sources has focused mostly on sources with strong social ties towards the active user, e.g. friends, and often claimed recommendations by personal sources with a strong tie strength to be superior to personalized recommendations provided by an impersonal source [30, 31, 1]. However, it remains unclear whether, and under which circumstances, the same applies to personal sources with weaker tie strengths. Only few attempts have been undertaken to compare personal sources to those generated by an impersonal RS (e.g. [18, 25, 5]).

When it comes to the accuracy of preference predictions, automated RS in general seem to perform better than users [18, 25]. However, this might be true only under certain circumstances. A study conducted by Krishnan et al. [18] showed that automated RS were more accurate in their recommendations, but only on average. Humans were found to be superior in recommending for preference profiles that lay outside the mainstream. A small number of expert users, who had a high knowledge of the item domain, was even able to outscore the automated system overall. While Krishnan et al. mainly focused on accuracy, they also suggest further dimensions for which it seems worthwhile to compare performances of humans and automated systems, namely diversity, novelty, and the trust a user has in the recommendations and recommendation source, respectively.

Trust is a multidimensional construct and a vast variety of terminology has been used to describe it. In this paper we follow the interdisciplinary model of trust in online situations[1] of McKnight et al. [22, 21]. Based on the well-established *theory of reasoned action* [10], this model explains mechanisms that lead to *trust-related behavior*, such as making an online purchase, sharing information with someone, or taking advice. Before trust-related behavior is performed, *trusting intentions* must be present. Trusting intentions directly describe the willingness and probability of a person to commit to a trust-related behavior. Three central constructs, which are interconnected among each other, are central for establishing trusting intentions: *disposition to trust*, *institution-based trust*, and *trusting beliefs*. The disposition to trust describes the general trusting stance and overall trustfulness of a person. A person's faith in humanity constitutes an example for disposition to trust. Institution-based trust, on the other hand, arises in certain situations and depends highly on the environment in which the trust-related behavior is going to take place (e.g. the trust in the Internet security during an online purchase). Finally, trusting beliefs concern characteristics of the trustee. According to McKnight et al. [21], *integrity* (the trustee's reliability and honesty), *benevolence* (favorable motives based on altruism, such as goodwill) and *competence* (the trustee's general ability to fulfill the truster's needs) constitute the trusting beliefs.

These characteristics align with other literature on trust, even though the terminology partly differs. In this sense, competence can be linked to a trustee's *expertise* [24, 34], which is often considered a key factor regarding trust towards an information source [13, 14]. Integrity, on the other hand, aligns with *trustworthiness* [24, 34]. Trustworthiness is sometimes also linked with benevolence, describing the willingness to share information and general goodwill of the trustee [21, 12]. Trusting beliefs can also concern a recommendation provider as special case of information source. Then, expertise describes the general ability of providing accurate recommendations, while trustworthiness focuses more on whether a recommendation provider is unbiased when creating recommendations [34].

[1] Note that this regards situations of *initial* trust only, i.e. when no long-term trust relationship has been established yet.

However, other aspects undoubtedly influence a user's trusting beliefs with respect to a recommendation source as well. One is the perceived *similarity*, also termed *homophily* [13] or *affinity* [14], between provider and receiver of recommendations [14, 34]. The effectiveness of perceived similarity is highly dependent on context and the situation in which a recommendation appears. Moreover, a higher similarity may sometimes lead to a decrease in the perceived trustworthiness [34]. In online situations, where the recommendation giver cannot be met in person, the degree of *perceived social presence* may affects a user's trust and reuse intention as well [6].

Several studies have investigated how such traits can effectively be conveyed by a RS. Herlocker et al. [15] found that increasing transparency, e.g. by explaining the reasons that led the system to recommend certain items, also increases the users' perceived trustworthiness of recommendations. More recently, Pu et al. [27] proposed a trust model for RS that aims to raise the perceived trustworthiness of predictions by explaining the advantages and trade-offs of each recommended item. Bonhard et al. [3] evaluated a movie recommender interface that visualized the recommendation giver's similarity to the target user in terms of shared interests and rating overlap. Their results show that users favored recommendations from people similar to themselves.

Besides, explicitly depicted traits of a recommendation source (especially expertise and similarity) seem to have seldom been subject of research regarding trust in RS.

METHOD

We conducted a user study consisting of two consecutive experiments under controlled conditions. Both experiments were based on a between-subject design. Under different conditions, distinct visual hints were provided regarding (1) the indication of the recommendation source itself, i.e. *personal* vs. *impersonal*, and (2) a personal recommendation source's degree of *expertise* and *similarity*. For all conditions we used the prototype described at the end of this chapter.

We recruited 88 participants (45 female) with an average age of 25.6 ($SD = 0.52$). Most of them were students (45) or employees (35). To interact with the RS prototype and to answer questionnaires, participants used a common web browser on a desktop PC (24" LCD-screen with a 1920×1200px resolution). In the following, we describe the experimental setup, prototype system, procedure as well as the questionnaires used in the study.

Initial Setup

Before beginning the first experiment, participants were asked to fill in an online questionnaire. This was used to evaluate prior knowledge of the domain using items from Knijnenburg et al. [16]. In addition, we assessed the participant's *disposition to trust* and *institution-based trust* using constructs by McKnight et al. [21]. We also asked participants to state general demographics.

Subsequently, participants were asked to complete a preliminary task that served for eliciting personal preferences. For this, 10 popular movies had to be rated on a 5-point scale. Movies were shown in random order and users could skip movies that were unknown to them. This is a common method known to yield good results [7] and that is frequently used in comparable user studies. Based on these ratings, user profiles were calculated according to [28].

First Experiment

Conditions: In the first controlled experiment, we wanted to compare the users' perception of personalized recommendations generated by personal and impersonal sources with respect to its influence on the perceived social presence (see **RQ1**). In particular, we defined four conditions with the following indications regarding the recommendation source:

- Conventional impersonal RS (RS^{auto})
- Personal source without further information ($User^{neutral}$)
- Personal source with indication of high expertise ($User^{exp}$)
- Personal source with indication of high similarity ($User^{sim}$)

Procedure: First, participants were randomly assigned to one of the conditions above. Participants in conditions $User^{neutral}$, $User^{exp}$, and $User^{sim}$ were told that their initially elicited preference were being sent to a specifically selected person sitting in the next room. Said person would then manually select and transmit back appropriate recommendations, which could be visualized afterwards. Even though our RS generated recommendations instantly in the background, the recommendation process was delayed artificially in order to simulate a manual selection process and thus introduce some realism. Based on the results of a pre-study, we limited the delay to 46 seconds, which was shown to be realistic while not appearing too tiresome. To add further realism, a loading screen was displayed during this waiting period, which consecutively depicted four processing steps: (1) *'Searching user.'*, (2) *'User received preferences.'*, (3) *'User selects items to recommend.'*, (4) *'Receiving recommendations.'*. Participants in condition RS^{auto} were presented with recommendations instantly.

Independently of the particular condition, 5 of the top-10 recommendations generated by a standard recommendation algorithm (explained later in this chapter) were shown to the user. Additionally, specific information about the recommendation source were depicted. For the three conditions with a personal recommendation source, one of two[2] iconic user avatars was shown randomly. In conditions $User^{exp}$ and $User^{sim}$ the avatar was combined with symbols indicating either high expertise or high similarity, respectively.

After participants had examined the recommendations and visual cues about the recommendation source closely, they were asked to rate recommendations and fill in a questionnaire. We used instruments by Gefen and Straub [11, 12] to measure the *perceived social presence*. The *tie strength* was assessed using constructs from Money et al. [23]. In addition, general recommender performance was measured using the *ResQue* evaluation framework introduced by Pu et al. [26]. In order to

[2]Even though we used gender-neutral avatars the depicted nicknames were not necessarily gender-neutral (*Maria463* and *Alexander721*). We assigned participants randomly to one of these two profiles to alleviate gender-based biases as much as possible.

measure trust-related effects caused by the variation of recommendation sources, we used constructs for *trusting beliefs* and *trusting intentions* by McKnight et al. [21].

Second Experiment

Conditions: In the second experiment, we focused solely on recommendations generated by a personal source. We wanted to inspect the influence of a sources' indicated expertise and similarity in more detail (see **RQ2**). Variation of both constructs yields, again, four conditions:

- Indication of high expertise and high similarity (*Exp*↑*Sim*↑)
- Indication of high expertise and low similarity (*Exp*↑*Sim*↓)
- Indication of low expertise and high similarity (*Exp*↓*Sim*↑)
- Indication of low expertise and low similarity (*Exp*↓*Sim*↓)

Procedure: After each participant was assigned randomly to a condition, the same waiting process was simulated as in the first experiment. Participants were again presented with the aforementioned cover story that their preferences were being shown to a selected person. Thereafter, the remaining 5 of the generated top-10 recommendations were displayed. This time visual hints for both characteristics of the recommendation source were shown. Similarly to the first experiment, participants were asked to rate the recommended items after examining them and the recommendation source thoroughly. The same questionnaires as in the first experiment were used here, albeit with minor changes. Specifically, we replaced constructs measuring the perceived social presence with items by Feick and Higie [9] evaluating *perceived expertise* and *similarity*.

Prototype

The same prototype was used in all experimental conditions. Its interface was designed to contain as little visual distractions as possible. Visual hints about recommendation source and its characteristics changed depending on the condition and served as independent variables in our setup. Recommendations for all conditions were generated using a standard Matrix Factorization (MF) [17] algorithm. We chose movie recommendations as our target domain. We used the MovieLens 20M dataset[3], which consists of 20 000 263 ratings for 27 278 movies provided by 138 493 users on a 5-point scale to calculate the underlying MF model. See Figure 1 for a screenshot of the prototype used during condition *Exp*↑*Sim*↓in the second experiment.

RESULTS

In the following, we present the results of our evaluation on how to influence user's initial trust by varying the aforementioned recommendation source's traits. First, we provide a general description of the sample's properties followed by an in-depth analysis of the two experiments conducted. In the last step, we utilize *Structural Equation Modeling* to reveal structural relationships between variations in the recommendation source's properties and trust-related constructs. We set the statistical significance level to $\alpha = .05$. For post hoc comparisons, we used Tukey-Test.

[3] http://grouplens.org/datasets/movielens/20m/

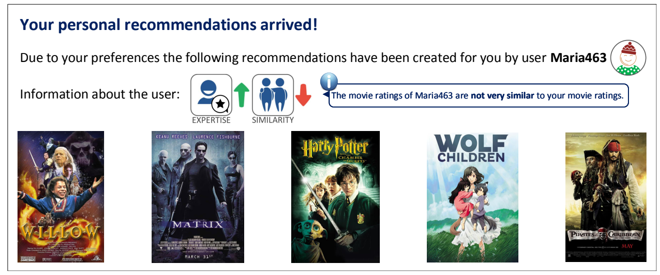

Figure 1. A screenshot of the used prototype. It shows a view during the second experiment in condition *Exp*↑*Sim*↓. Currently a tooltip is shown for the similarity.

General

Overall, participants reported that they had an average knowledge about movies ($M = 3.11$, $SD = 1.1$) and a rather high trust in technology ($M = 3.85$, $SD = 0.54$). The general *disposition to trust* another person was rated relatively high ($M = 4.64$, $SD = 1.24$) while *institution-based trust* was rated moderately ($M = 3.45$, $SD = 0.93$).

First Experiment

Checking for differences between groups, we found a weak indication of sample selection bias, $F(3,84) = 2.65, p = .05, \eta_p^2 = .09$. Post hoc tests revealed that participants in condition $User^{sim}$ had a significantly lower dispositional trust than the ones in $User^{exp}$, $p = .041$. As a result, in order to maintain internal validity, the follow-up analyses were carefully controlled for the effect of the confounding variable *disposition to trust*. We found no significant differences for the other trust-related antecedent *institution-based trust*, $F(3,84) = 1.13, p = .342, \eta_p^2 = .04$. We conducted a one-way MANCOVA, with *disposition to trust* as a covariate, to compare the effect of personal vs. impersonal recommendation (i.e. *condition*) sources on non-static trust-related constructs, i.e. *trusting beliefs* and *trusting intentions*, as well as *perceived social presence* and *use intentions*. The estimated marginal means for the four conditions are presented in Table 1. The multivariate effect with $F(12,246) = 1.26, p = .421, \eta_p^2 = .05$ for *condition* was not significant. However, we concluded that further explorative analyses could lead to meaningful insights. Analyzing the between-subject effects, we could observe that *condition* has a small effect on *social presence*, $F(3,83) = 2.59, p = .058, \eta_p^2 = .086$. Although the covariate's overall effect was not significant, $F(4,80) = 1.5, p = .209, \eta_p^2 = .07$, inspection of the between-subject effects yielded a significant influence on *trusting intentions*, $F(1,83) = 4.96, p = .029, \eta_p^2 = .06$. On average, participants rated the quality of each of their recommendations with $M = 2.9$ ($SD = 0.85$) on a 5-point scale. No differences between groups could be identified.

Second Experiment

For the second experiment, we used one-way MANOVA to test the effect from *condition*, i.e. the variation of *similarity* and *expertise* of a personal recommendation source, on *perceived expertise*, *perceived similarity*, *trusting belief*, *trusting intentions* and *use intentions*. See Table 2 for descriptive statistics. The multivariate effect for *condition* was significant, $F(15,246) =$

Table 1. Estimated Marginal Means and Standard Error for all dependent variables surveyed during the first experiment.

	Impersonal		Personal					
	RS^auto		*User^neutral*		*User^exp*		*User^sim*	
Construct	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Social presence	2.81	0.25	3.48	0.25	3.1	0.26	2.52	0.26
Trusting beliefs	4.51	0.28	4.47	0.28	4.61	0.28	3.85	0.28
Trust. intentions	4.06	0.25	4.04	0.25	3.97	0.25	3.45	0.25
Use intentions	3.17	0.21	3.37	0.21	3.37	0.21	2.85	0.21

$2.22, p = .006, \eta_p^2 = .12$. Inspecting the univariate ANOVAS, we could observe significant influences on *perceived expertise*, $F(3,84) = 4.44, p = .006, \eta_p^2 = .14$, and a tendency for *perceived similarity*, $F(3,84) = 2.01, p = .12, \eta_p^2 = .07$. Post-Hoc tests indicate that *perceived expertise* was rated significantly higher for *Exp↑Sim↑*than for *Exp↓Sim↓*($p = .006$) as well as *Exp↓Sim↑*($p = 0.27$). Comparable to the results of the first experiment, the average rating for each recommendation was assessed with $M = 2.9$ ($SD = 0.93$). Again, there were no differences between groups.

Table 2. Mean Values and Standard Deviations for all dependent variables surveyed during the second experiment.

	Exp↓Sim↓		*Exp↑Sim↓*		*Exp↓Sim↑*		*Exp↑Sim↑*	
Construct	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Perc. Expertise	2.76	1.01	3.17	1.19	2.91	0.87	3.72	0.59
Perc. Similarity	2.09	0.92	2.2	0.96	2.71	1.08	2.54	0.89
Trusting Beliefs	3.86	1.76	4.3	1.72	4.24	1.3	4.38	1.23
Trust. Intentions	3.29	1.41	3.54	1.6	4.06	1.43	3.9	1.33
Use Intentions	2.98	1.3	2.8	1.1	3.24	1.2	3.2	0.9

Structural Equation Modeling

Since we were particularly interested in deriving suggestions about how and which properties of the recommendation source influences trust into the recommendations (see **RQ3**), we used *Structural Equation Modeling* (SEM) to further investigate the effect that the degree of similarity and expertise have on trust-related constructs.

Our theoretical model is displayed in Figure 2. It yielded a good fit with the data ($\chi^2(106) = 126.059$, $p = .089$, $CFI = .982$, $TLI = .977$, $RMSEA = 0.046$). The model indicates that displaying whether a recommendation source is an expert or not (*condition*) significantly influences *perceived expertise*. However, with $R^2 = .10$ the amount of explained variance is rather low. Similarly, *perceived similarity* towards the recommendation source is influenced by indicating similarity in the interface (*condition*). *Disposition to trust* additionally significantly predicts *perceived similarity*. Together, *condition* and *disposition to trust* account for 20% of the variance in *perceived similarity* ($R^2 = .20$).

Both *perceived expertise* and *perceived similarity* are related to *trusting beliefs* with $R^2 = .73$. While there is no direct effect from *disposition to trust* toward *trusting beliefs* ($\beta = 0.06$, $p = .416$), there is a significant indirection via *perceived similarity* ($\beta = 0.08$, $p = .05$) resulting in a total effect of $\beta = .13$, $p = .054$.

Trusting intentions is directly influenced by *trusting beliefs* and *perceived similarity*. Some portion of the predictive power from *perceived similarity*, however, gets mediated by *trusting beliefs* yielding an indirect effect of $\beta = 0.187$, $p = .012$ and a total effect of $\beta = 0.445$, $p < .001$. Another mediation is present for *perceived expertise* via *trusting beliefs* ($\beta = 0.501$, $p < .001$). Combined with the non-significant direct relation ($\beta = -0.118$, $p = .292$), the model achieves a total effect of $\beta = 0.384$, $p = .001$. Finally, one can observe a significant total effect ($\beta = 0.189$, $p = .01$) from *disposition to trust* on *trusting intentions* when combining the non-significant direct ($\beta = 0.132$, $p = .052$) and indirect ($\beta = 0.06$, $p = .055$) influence via the route *perceived similarity* and *trusting beliefs*. Put together, all influencing variables yield an amount of explained variance of $R^2 = .78$.

DISCUSSION

The SEM (Figure 2) shows that the conditions during the second experiment had a significant influence on *perceived expertise* and *perceived similarity* of the recommendation source. This is especially noteworthy given that the conditions only featured minor visual differences: small arrows next to an icon for each trait. The implication is that these subtle indications were indeed persuasive. Following the paths of our SEM further, one can observe that *perceived expertise* and *perceived similarity* influenced the user's *trusting beliefs* and *trusting intentions*. Thus, it seems theoretically possible to control the user's trust in recommendations by only manipulating the depicted expertise and similarity of a recommendation source. This also seems to hold for the *perceived social presence* during the first experiment. Here the condition had an influence on the *perceived social presence*, even though the measured effect was rather low.

With respect to our initial research questions, our findings seem promising. The depiction of an iconic user avatar influenced the perceived social presence (**RQ1**). Moreover, providing indications of expertise and similarity had an impact on the perception of these traits (**RQ2**). Finally, these traits were shown to influence the users' trust (**RQ3**).

As stated above, the *condition* comparing indicators of expertise and similarity had influence on *perceived expertise* and *perceived similarity*. However, *condition* only accounts for roughly 10% of variance in both. We ascribe this to the fact that users in our setting had no real explanation of these traits other than simple icons with labels. It follows that the main portion of variance was caused by factors not identified in our model.

While *perceived expertise* and *perceived similarity*, taken together, have a notable influence on *trusting beliefs*, *perceived similarity* also directly affects *trusting intentions*. If *trusting intentions* is composed of the willingness to share information [21], the direct relation can be explained as follows: One is more likely to share information with a seemingly similar person than with an expert. *Trusting intentions* thus appear to be more reciprocal than *trusting beliefs*.

Since the mean values for *perceived similarity* are, overall, lower than for expertise (see Table 2), we assume that it's

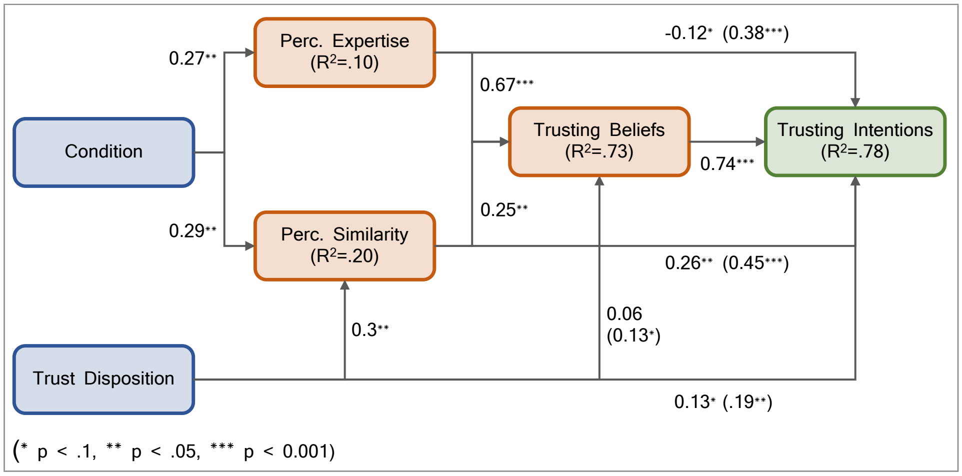

Figure 2. Structural Equation Model for Experiment 2. Independent (exogenous) variables are marked in blue, moderators in orange, dependent (endogenous) variables in green. We report corrected β-weights for direct effects. If present, corrected β-weights for total effects are reported in brackets.

harder to convey similarity. This is further underlined by the fact that, even though we found an influence of *condition* to *perceived similarity* in the SEM, there was no significant correlation in ANOVA. We conclude that the measured *perceived similarity* is a borderline case: While users indeed perceived some degree of similarity depending on the visual indications, they need more information about a recommendation source than we presented to really decide if there exists an overlap in interests with respect to the target domain. Interestingly, our definition of similarity, which is based on rating behavior (see tooltip in Figure 1), aligns with often-used phrases used for providing explanations in automated recommender systems, e.g. "Similar users also bought...". Considering our findings, we question whether phrases like these are effective enough means to convey similarity in a plausible way. It might thus be beneficial to explore the influence of richer visual elements like, for instance, those found by Bonhard and Sasse [4].

When examining descriptive values in Table 1, the highest average *perceived social presence* can be found in condition $User^{neutral}$. Both other conditions of personal sources are lower in tendency, with the condition $User^{sim}$ also showing significant differences. This is surprising because similar visual indicators for the source were used. The trend continues when examining other mean values of the condition $User^{sim}$. Strangely, all values were lower in their tendency compared with other conditions. Even though we do not have a final explanation for this effect, we assume it is the result of some form of bias that we were unable to measure. (Note that we identified *disposition to trust* as a confounding variable, which we subsequently integrated as a covariate into our analyses.)

CONCLUSIONS AND OUTLOOK

In this paper we have introduced a user study to compare personal and impersonal recommendation sources and the influences of traits of personal recommendation sources on a user's trust in recommendations. When comparing distinct personal sources with different traits, we found that the recommendation provider's expertise and similarity to the target user do influence the latter's trust in the recommended items. We further investigated how traits of a personal recommendation source can be conveyed visually. In particular, we showed that very subtle visual cues are sufficient to trigger perceptions of social presence, expertise and similarity towards a recommendation source. In addition, we add evidence that perceptions of these traits can increase a user's trust in recommendations.

These findings yield several interesting implications. In systems that already exploit personal recommendation sources, subtle indications of expertise and similarity might be considered for increasing the user's trust in the generated predictions. In such systems, these traits may also serve as criteria for deciding which peers to select as candidates for providing recommendations. In systems that do not rely on personal recommendation sources, including elements for conveying social features could be considered. This is due to our findings that social presence alone may increase trust in the recommendations.

Our approach has some limitations. The effects of highlighting properties regarding the recommendation source on the users' perceptions have been of minor significance. This might be due to the very subtle elements we used. Nonetheless, the perceived expertise and similarity clearly have a significant influence on building short-term trust. In order to further increase this influence, additional trust-building traits of the recommendation source could be depicted visually. Here, related literature indicates traits such as the source's benevolence or its integrity as possible candidates. However, depicting these traits might not be as straightforward and could require the development of more sophisticated visual elements. For the future, we are thus interested in further investigating means to communicate properties of recommendation sources effectively. We also plan to integrate them with automated recommender functions, which may lead to interesting reciprocal effects.

REFERENCES

1. Sam Banks, Rachael Rafter, and Barry Smyth. 2015. The Recommendation Game: Using a Game-with-a-Purpose

to Generate Recommendation Data. In *Proceedings of the 9th ACM Conference on Recommender Systems (RecSys '15)*. ACM, New York, NY, USA, 305–308.

2. Shlomo Berkovsky, Ronnie Taib, and Dan Conway. 2017. How to Recommend?: User Trust Factors in Movie Recommender Systems. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion (IUI '17 Companion)*. ACM, New York, NY, USA, 287–300. DOI: http://dx.doi.org/10.1145/3025171.3025209
3. Philip Bonhard, Clare Harries, John McCarthy, and M. Angela Sasse. 2006. Accounting for Taste: Using Profile Similarity to Improve Recommender Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 1057–1066. DOI: http://dx.doi.org/10.1145/1124772.1124930
4. P. Bonhard and M. A. Sasse. 2006. 'Knowing Me, Knowing You' – Using Profiles and Social Networking to Improve Recommender Systems. *BT Technology Journal* 24, 3 (2006), 84–98. DOI: http://dx.doi.org/10.1007/s10550-006-0080-3
5. Shuo Chang, F. Harper, Lingfei He, and Loren Terveen. 2016. CrowdLens: Experimenting with Crowd-Powered Recommendation and Explanation. In *Proceedings of the Tenth International AAAI Converence on Web and Social Media (ICWSM)*. 52–61.
6. Jaewon Choi, Hong Joo Lee, and Yong Cheol Kim. 2009. The influence of social presence on evaluating personalized recommender systems. In *Pacific Asia Conference on Information Systems (PACIS)*.
7. Paolo Cremonesi, Franca Garzottto, and Roberto Turrin. 2012. User Effort vs. Accuracy in Rating-based Elicitation. In *Proceedings of the Sixth ACM Conference on Recommender Systems (RecSys '12)*. ACM, New York, NY, USA, 27–34. DOI: http://dx.doi.org/10.1145/2365952.2365963
8. Dale F. Duhan, Scott D. Johnson, James B. Wilcox, and Gilbert D. Harrell. 1997. Influences on consumer use of word-of-mouth recommendation sources. *Journal of the Academy of Marketing Science* 25, 4 (1997), 283. DOI: http://dx.doi.org/10.1177/0092070397254001
9. Lawrence Feick and Robin A. Higie. 1992. The Effects of Preference Heterogeneity and Source Characteristics on Ad Processing and Judgements about Endorsers. *Journal of Advertising* 21, 2 (1992), 9–24. DOI: http://dx.doi.org/10.1080/00913367.1992.10673364
10. Martin Fishbein and Icek Ajzen. 1975. *Belief, attitude, intention and behavior: An introduction to theory and research.* Addison-Wesley, Reading, Mass. [u.a.].
11. David Gefen and Detmar W. Straub. 1997. Gender Differences in the Perception and Use of E-Mail: An Extension to the Technology Acceptance Model. *MIS Quarterly* 21, 4 (1997), 389–400.
12. David Gefen and Detmar W. Straub. 2004. Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. *Omega* 32, 6 (2004), 407–424. DOI: http://dx.doi.org/10.1016/j.omega.2004.01.006
13. Mary C. Gilly, John L. Graham, Mary Finley Wolfinbarger, and Laura J. Yale. 1998. A Dyadic Study of Interpersonal Information Search. *Journal of the Academy of Marketing Science* 26, 2 (1998), 83–100. DOI:http://dx.doi.org/10.1177/0092070398262001
14. Thomas Heath, Enrico Motta, and Marian Petre. 2006. Person to person trust factors in word of mouth recommendation. In *Proceedings of the 2006 Conference on Human Factors in Computing Systems (CHI '06)*.
15. Jonathan L. Herlocker, Joseph A. Konstan, and John Riedl. 2000. Explaining Collaborative Filtering Recommendations. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work (CSCW '00)*. ACM, New York, NY, USA, 241–250. DOI: http://dx.doi.org/10.1145/358916.358995
16. Bart P. Knijnenburg, Martijn C. Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. Explaining the User Experience of Recommender Systems. *User Modeling and User-Adapted Interaction* 22, 4-5 (2012), 441–504.
17. Yehuda Koren, Robert M. Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. *IEEE Computer* 42, 8 (2009), 30–37. DOI: http://dx.doi.org/10.1109/MC.2009.263
18. Vinod Krishnan, Pradeep Kumar Narayanashetty, Mukesh Nathan, Richard T. Davies, and Joseph A. Konstan. 2008. Who Predicts Better?: Results from an Online Study Comparing Humans and an Online Recommender System. In *Proceedings of the 2008 ACM Conference on Recommender Systems (RecSys '08)*. ACM, New York, NY, USA, 211–218.
19. Feng Li and Timon Du C. 2011. Who is talking? An ontology-based opinion leader identification framework for word-of-mouth marketing in online social blogs. *Decision Support Systems* 51, 1 (2011), 190–197. DOI: http://dx.doi.org/10.1016/j.dss.2010.12.007
20. Roger C. Mayer, James H. Davis, and F. David Schoorman. 1995. An Integrative Model of Organizational Trust. *The Academy of Management Review* 20, 3 (1995), 709–734.
21. D. Harrison McKnight, Vivek Choudhury, and Charles Kacmar. 2002. Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research* 13, 3 (2002), 334–359. DOI:http://dx.doi.org/10.1287/isre.13.3.334.81
22. D. Harrison McKnight, Larry L. Cummings, and Norman L. Chervany. 1998. Initial Trust Formation in New Organizational Relationships. *The Academy of Management Review* 23, 3 (1998), 473–490.

23. R. Bruce Money, Mary C. Gilly, and John L. Graham. 1998. Explorations of National Culture and Word-of-Mouth Referral Behavior in the Purchase of Industrial Services in the United States and Japan. *Journal of Marketing* 62, 4 (1998), 76–87.

24. Daniel J. O'Keefe. 1990. *Persuasion: Theory and research*. Sage Publications, Inc, Thousand Oaks, CA, US.

25. Peter Organisciak, Jaime Teevan, Susan Dumais, Robert C. Miller, and Adam Tauman Kalai. 2014. A Crowd of Your Own: Crowdsourcing for On-Demand Personalization. In *Second AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2014)*. AAAI - Association for the Advancement of Artificial Intelligence.

26. Pearl Pu, Li Chen, and Rong Hu. 2011. A User-centric Evaluation Framework for Recommender Systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems (RecSys '11)*. ACM, New York, NY, USA, 157–164.

27. Pearl Pu, Li Chen, and Rong Hu. 2012. Evaluating recommender systems from the user's perspective: survey of the state of the art. *User Modeling and User-Adapted Interaction* 22, 4 (2012), 317–355. DOI: http://dx.doi.org/10.1007/s11257-011-9115-7

28. Steffen Rendle and Lars Schmidt-Thieme. 2008. Online-updating regularized kernel matrix factorization models for large-scale recommender systems. In *Proceedings of the 2008 ACM conference on Recommender systems*. ACM, 251–258.

29. Sylvain Senecal and Jacques Nantel. 2004. The influence of online product recommendations on consumers' online choices. *Journal of Retailing* 80, 2 (2004), 159–169.

30. Rashmi Sinha and Kirsten Swearingen. 2001. Comparing Recommendations Made by Online Systems and Friends. In *In Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries*.

31. Donnavieve Smith, Satya Menon, and K. Sivakumar. 2005. Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of Interactive Marketing* 19, 3 (2005), 15–37. DOI: http://dx.doi.org/10.1002/dir.20041

32. Patricia Victor, Martine de Cock, and Chris Cornelis. 2011. Trust and Recommendations. In *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor (Eds.). Springer US, 645–675. DOI: http://dx.doi.org/10.1007/978-0-387-85820-3_20

33. Kyung-Hyan Yoo and Ulrike Gretzel. 2008. The influence of perceived credibility on preferences for recommender systems as sources of advice. *Information Technology & Tourism* 10, 2 (2008), 133–146. DOI: http://dx.doi.org/10.3727/109830508784913059

34. Kyung-Hyan Yoo, Ulrike Gretzel, and Markus Zanker. 2015. Source Factors in Recommender System Credibility Evaluation: Recommender Systems Handbook. In *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). Springer US, Boston, MA, 689–714. DOI: http://dx.doi.org/10.1007/978-1-4899-7637-6_20