A Diary Study of Social Explanations for Recommendations in Daily Life

Zhirun Zhang
zrzhang@comp.hkbu.edu.hk
Department of Computer Science,
Hong Kong Baptist University
Hong Kong, China

Yucheng Jin
yuchengjin@hkbu.edu.hk
Department of Computer Science,
Hong Kong Baptist University
Hong Kong, China

Li Chen
lichen@comp.hkbu.edu.hk
Department of Computer Science,
Hong Kong Baptist University
Hong Kong, China

ABSTRACT

We report a diary study of the explanations for the recommendations to characterize the social features in these explanations recorded by five participants over two months. The study reveals several social explanation categories (e.g., personal opinions and personal experiences) and their relationship with user contexts (e.g., location, relevant experience) and recommender attributes (e.g., integrity, expertise) illustrated in a network diagram. Specifically, personal opinions and experiences are two prominent social explanations, mainly associated with user contexts (e.g., users' preferences and users' experiences) and several recommender attributes (e.g., politeness, benevolence, and experience). Finally, we discuss several design implications for social explanations and anticipate the value of our findings regarding designing personalized social explanations in recommender systems that aim to build rapport with users, such as conversational recommender systems.

CCS CONCEPTS

• **Human-centered computing → User centered design**; *Field studies*; • **Information systems → Recommender systems**.

KEYWORDS

Social explanations, recommender systems, diary study

ACM Reference Format:
Zhirun Zhang, Yucheng Jin, and Li Chen. 2022. A Diary Study of Social Explanations for Recommendations in Daily Life. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '22 Adjunct), July 4–7, 2022, Barcelona, Spain.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3511047.3537681

1 INTRODUCTION

Users' preferences for products or any items tend to be affected by the actions of the people surrounding them [7]. For example, due to social trust, users are more likely to watch a new movie if it is recommended by their friends [34]. Thus, several researchers proposed social recommender systems that incorporate the social network information (e.g., friendship and common interests) in recommendation algorithms to improve the quality and user acceptance of recommendations [11, 27, 31]. For example, replacing rating-similarity neighborhoods with the social relationship in collaborative filtering could lead to better performance in taste-related domains such as music and movie recommendations [9].

Moreover, the underlying social information of the recommendation algorithm can be used to explain the recommendations, known as social explanations, which could improve the perceived quality of recommendations and support decision making [28]. The existing studies have identified what types of social explanations could be [21, 22] and how they improve the user experience of recommender systems [10, 22]. However, little is known about how users feel about social explanations and whether their attitudes (e.g., good or bad) depend on other factors such as user contexts and recommender attributes. Therefore, we might ask *what user contexts and recommender attributes are particularly associated with different social explanations?*

To fill this research vacancy, we conducted a two-month diary study with five participants to understand their perception of the social explanations for the recommendations they encountered in their daily lives. We then performed a thematic analysis for the diary data from three aspects: explanations, user contexts, and recommender attributes, which are determined based on the framework of explaining the user experience of recommender systems [15]. By coding the explanations, we identify four categories of social explanations: relation-based, third-party opinions, personal opinions, and personal experiences. Furthermore, according to participants' feedback, we analyzed how good or poor social explanations relate to user contexts (e.g., location, relevant experience) and recommender attributes (e.g., integrity, expertise). We think our findings shed light on designing personalized social explanations for recommendations.

2 RELATED WORK

Social explanations could be presented visually or textually in recommender systems. Visualizing social information behind recommendations is an effective way to improve the transparency and explainability of social recommender systems [33]. Tsai and Brusilovsky [33] investigated various visualizations (e.g., Venn diagrams, maps, radar charts) to explain social recommendations, including interest similarity, topic similarity, co-authorship similarity, etc. Besides, TasteWeights is an interactive visualization that shows the links among user profiles, context, and music recommendations, which leads to higher user satisfaction [3]. Despite the advantages of visualizations in explaining social recommendations, visualizations need more display space than texts. Besides, interacting with

UMAP '22 Adjunct, July 4–7, 2022, Barcelona, Spain

ACM ISBN 978-1-4503-9232-7/22/07…$15.00
https://doi.org/10.1145/3511047.3537681

visualization-based explanations relies on graphical user interfaces (GUIs). In contrast, textual explanations are more adaptive to other user interfaces, such as conversational user interfaces (CUIs). Thus, this study focuses on text-based social explanations and explores how user contexts and recommender attributes influence users' feelings about these explanations.

A prior study identified four categories of textual social explanations based on the content of a movie discussion data set, including item features, third-party opinions, personal opinions, and personal experiences [22]. Papadimitriou et al. [21] revealed explanation styles for social recommender systems based on three essential resources for explanations: similar users, similar items, and item features. A previous study compared different styles of explanation and found that socio-centric explanations were less persuasive than content-based explanations [16]. However, Hayati et al. [10] showed that several sociable strategies, such as personal opinions, personal experience, and user similarity, have a significant effect on the success of recommendations. Therefore, we are curious about what kind of social explanation might help the user make an informed decision on recommendations and what factors might influence user attitude about social explanations.

Previous work indicates that human factors might affects the user perception of recommendation and its explanations. Szymanski et al. [32] investigated the effect of domain knowledge on visual, textual, and hybrid explanations. The results showed that different expertise groups had a different understanding of visual explanations. Besides, Kouki et al. [16] found that personality affected the preference for explanation styles and the number of explanation styles. Millecamp et al. [19] explored the impact of various personal characteristics on user perception and interaction in a music explainable recommender system and found that the need for cognition influenced the user's confidence when using the visualization-based explanations. Moreover, the effect of explanations on situation awareness was influenced by different levels of self-reported task familiarity [26]. These studies indicate the factors influencing user perception of explanations, which help us determine what factors should be considered while looking at social explanations.

3 DIARY STUDY

We chose a diary study, a method that asks participants to record their daily activities and experience in a period of time [24, 30], to understand user perception of social explanations they have experienced in daily life. The method has been used in the area of human-computer interaction, such as [6, 30, 35]. Compared with laboratory-based approaches, diary study is recorded freely under different actual environments, which influence users' perception [29]. In other words, the user experience recorded in the diary study is closer to users' actual experience than in laboratory-based approaches. Moreover, as explanations of recommendations in daily life are diverse and ephemeral, daily records in the diary study could reduce biases by retrospect [29].

Table 1: Participants' demographics in the diary study

Participant	**Gender**	**Age range**	**Education level**
P1	F	35-44	PhD
P2	M	18-24	Bachelor
P3	F	18-24	Bachelor
P4	F	18-24	Bachelor
P5	F	18-24	Master

3.1 Participants and Procedure

Because the quality of a diary study mainly depends on the quality of participants' diary entries, it is more challenging to train participants compared with interviews or usability study [29]. Therefore, we recruited five researchers (see Table 1) in recommender systems to participate in this diary study. We believe their expertise in recommendations allows them to observe the explanations for recommendations in their daily lives and make a high-quality description of observation and perception.

We first asked participants to record the explanations in their daily life by the instruction below:

> The purpose of this study is to record the good/bad explanations we have seen or heard in our daily life, for example, from your friends or other people you know, or when you go to a store, a restaurant, a travel agency, or any websites/apps that may provide any service, information, or item recommendations for assisting your decision-making. Please record the explanation that has impressed you, which might be *good* in terms of helping you make a more informed decision and build a trust relationship with the recommender, or *bad* if you have any negative feelings after you receive such an explanation. Moreover, you may explicitly request "recommendation" and ask "why" in some situations you think appropriate.

Besides the explanation, participants are also asked to write down their personal opinions, describe the scenario when receiving the explanation, explain why they think the explanation is good or bad, and take a screenshot of the interface of the explanation if possible (see Table 2). Each participant recorded about one explanation every day. Two months later, we ceased the study as participants responded that they could not find any new explanations with a different pattern from the previously recorded explanations. Finally, we recorded 146 entries in the diary study and dropped seven entries because of low relevance to recommendations (5), lack of explanations (1), or not from a user perspective (1).

3.2 Analysis

Focusing on the factors that influence user perception in the diary entries, the first author generated codes and developed a codebook with the third author using thematic analysis [4, 25]. After that, the first author and a research assistant coded the diary entries following the guidelines of qualitative research [20]. They coded the same subset of data (15 entries) independently using the codebook, computed the reliability of each code, discussed the inconsistencies,

Table 2: An Example recorded in our Diary Study

Component	Feedback from the participant
User	P1
Date	6/17/ 2019
Explanation	From a friend: "I recently watched a movie. It is the best one I have watched this year. It is really good."
Interface sample (if any)	N/A
Tags on this explanation (your personal opinions)	Not personalized, because this person does not know my movie preferences, but as she gave such positive review, I am curious to watch this recommended movie.
Context/scenario (in what conditions you get this explanation)	I got this recommendation when we took lunch together. She said this movie was just played in HK for four days, and it is really worth watching it in her opinion.
Remark (why you think it is good or bad)	Neutral: 1. I am interested in watching this movie, but because the recommender does not have common interests with me, I am a little hesitant whether I will like it as her. 2. Good recommendation, but the final opinion may depend on my watching experiences. If I do not like it, probably I will not accept this person's movie recommendation in the future.

and refined the codebook. The procedures were repeated on different subsets of data until the overall interrater reliability was strong, i.e., the average Cohen's Kappa of all codes is greater than 0.8 [1]. The Cohen's Kappa of all codes ranges from 0.59 to 1 (Average = 0.88). Finally, the first coder followed the refined codebook to code the remaining data set.

After coding, we filtered out 29 explanations that do not contain social features represented by our defined social explanation categories (see table 3). Eventually, we got 110 diary entries that contain social explanations in the recording for further analysis. We illustrate the relationship between different factors that contribute to the user's attitude towards different social explanations in a network diagram. In the following text, we will use C and P to respectively represent the diary entry and the participant when quoting the participant's diary entry.

4 RESULTS AND DISCUSSION

In the 110 diary entries, 78.18% (N=86) of them were from humans, and 21.82% (N=24) were from systems. According to our defined explanation categories, 7.27% (N=8) mentioned relation-based explanations, 21.81% (N=24) mentioned third-party opinions, 80.00% (N=88) mentioned personal opinions, and 40.00% (N=44) mentioned personal experiences. Regarding the participants' feelings, 61.82% (N=68) recorded "good" explanations, 29.09% (N=32) recorded "bad" explanations, and 9.09% (N=10) recorded "neutral" explanations, which contained positive and negative feelings. The recommendations in the diary study covered various domains: including entertainment (e.g., games, music, and movie), education (e.g., course, book, and article), service (e.g., restaurant, hotel, and tourist sight), electronics (e.g., camera, smartwatch, and software), living supplies (e.g., cup, shampoo, and light), and beauty (clothes, shoes, and makeup), etc.

[1] 0.01–0.20 as none to slight, 0.21–0.40 as minimal, 0.41– 0.59 as weak, 0.60–0.79 as moderate, 0.80–0.90 as strong, and > 0.90 as almost perfect agreement [18].

4.1 Codebook

The codebook consists of three parts: explanations, user contexts, and recommender attributes. We develop a classification for each part based on the existing research definitions and our observations.

4.1.1 Explanations. Based on the social features represented in explanations, we categorized social explanations into *relation-based*, *third-party opinions*, *personal opinions*, and *personal experiences*. Table 3 shows the definitions and examples of each explanation category. Relation-based explanations, also known as collaborative explanations, can be made based on explicit or implicit relationships, which are commonly used to build social recommender systems [12]. The opinions about the recommended items can be either from the recommender (e.g., a human or a system) or a third party (who has no relationship with the user), while the experiences shared in the explanations are often based on human recommenders' experiences [22]. The relation-based explanations are based on the explicit relationship (e.g., friends, classmates) or implicit relationship (e.g., sharing similar preferences). In contrast, third-party opinions are provided by a person or an agency that has no relationship with the user. For example, a mobile phone is recommended based on some mobile phone experts' opinions. During the coding, we also identified other types of explanations in the data, but we excluded them because they were irrelevant to social explanations.

4.1.2 User Contexts. We also identified different user context categories, which can be categorized into users' external situations (External), users' internal states (Internal), and the relationship between users and recommenders (Relation) (see Table 4).

Users' external situations (External). The external situation/scenario can influence the user's attitude toward the explanation. Three participants showed that the user's location (nine occurrences) influences their decisions. For example, the explanation did not consider the user's location may make the user hesitate, "*Both of my friends live in mainland China, but I live in HK. I doubt the availability of applet in HK.*" (P5, C131). Moreover, three participants considered the environmental condition (four occurrences), including weather

Table 3: The classification of social explanations in the codebook

Category	Description	Example
Relation-based (N=8)	The item is recommended or from a person who has explicit relationship with the user *(explicit relationship)*	"*Read by xx friend(s)*" (P2, C92)
	The item is recommended based on a person who has similar preference but no explicit relationship with the user *(implicit relationship)*	"*The following are recommended by users of this website*" (P2, C74)
Third-party opinions (N=20)	The opinions are from someone who is not the recommender and has no relationship with user	"*A customer said, the food tastes good and I will come again*" (P2, C80)
Personal opinions (N=89)	The subjective opinion from the recommender, such as feeling and evaluation	"*I think it is not very sweet and has a very good taste.*"(P3, C10)
Personal experiences (N=44)	Human recommenders have experienced the item and make recommendations based on their experiences	"*When I learn a new language or a new algorithm, I use this platform to practice*" (P5, C58)

Table 4: The classification of user contexts in the codebook

Category	Description	Occurrences
Users' external situations (External)		
Environmental Condition	The condition surrounding users, such as weather and atmosphere	4
Companion	Such as being alone, with close friend	2
Time	The time/schedule of the user	3
Location	The location of the user	9
Users' internal states (Internal)		
Demographics	The demographic information, such as gender and age	6
Personality	Such as openness to experience	4
Physical	The body or the feeling on the body, such as thirst and fatigue	6
Domain Knowledge	The user has domain knowledge of the item	2
Relevant Experience	Whether the user has relevant experience of an item which is similar to the recommendation	15
Emotion	A feeling such as happiness, anger, or sadness	2
Target Audience	Seeking items for themselves or for others	3
Preference	Whether the user has a strong preference	25
Relationship between users and recommenders (Relation)		
Trusted Relationship	The user trusted the recommender before receiving the recomemndation	10
Common Interest (human)	The recommender has common interests with the user	11
Successful Experience	The recommender has successfully recommended items to the user	9

and atmosphere. For example, P3 received the explanation that the recommended fan made the recommender cold and can be used in summer; she liked the explanation because"*It is suitable for summer seasons*" (P3, C46). Another example is the atmosphere, "*Friendly and cheerful atmosphere makes it easier to accept the recommendations" (P4, C5)*. Two participants mentioned the time (three occurrences). For example, P2 stated, "*At that time, I did not want to learn a new skill of singing whistles because I had other places to go.* " when the recommender said "*I can teach you how to use the whistle.*" as an explanation (P2, C127). In addition, two participants considered the companion (two occurrences), including with friends and with the child, e.g., "*I intend to take my friends to that place if we decide to have an annual meeting in Shanwei next time.*" because the explanation mentioned the friends gathering (P3, C47).

Users' internal states (Internal). we regarded users' physical or psychological states in diary entries as the user's "internal" context. Four participants considered the demographic information (six occurrences) in the decision-making process, including gender, age, and culture. For example, when receiving an explanation that "*becoming an exquisite stylish girl instantly.*", P1 felt that "*the explanation is useless when the recommendation is not accurate.*" (P1, C52). The mismatched gender in the explanation directly led the participant to determine that the recommendation was inappropriate, although the recommended item, a makeup magazine, could be read by the male. Another example is related to culture. When a person speaking Cantonese recommended a TV show and explained that "*You can learn Cantonese from it.*" to P5, who cannot speak Cantonese, P5 stated, "*My friend knows that I am learning Cantonese and she teaches me sometimes. So her recommendation is persuasive.*" (P5, C144). Moreover, although the diary entries did not explicitly reflect users' personalities, we inferred that the participants who are open to new things might tend to accept the recommendation (four occurrences). For instance, P1 stated, "*I am curious to watch it given his very positive comments on it.*" (P1, C8). Furthermore, four participants reflected that physical status (six occurrences), including the feeling of the body (thirst and fatigue) and the body condition (skin condition and physical fitness), changed their attitudes. For example, P4 wrote, "*Although the chocolate flavor is usually good, at that time I was thirsty, so it is not a suitable choice*" (P4, C20). Two participants used their domain knowledge (two occurrences) to judge the explanation. For example, P2 disagreed with the explanation, "*As an expert in this field, I know machines cannot beat humans at the current stage.*" (P2, C138). Four participants' feelings were affected when they had relevant experience (15 occurrences). The relevant experience may increase users' interest in the recommended items. For example, P5 wrote, "*The brand he recommends is famous for candy which I often ate in my childhood, so I am interested in.*", which was emphasized in the explanation "*for the childhood memory*" (P5, C66). Moreover, one participant mentioned emotion (two occurrences) that influences the attitude, including happiness and sadness, for example, "*I accepted it maybe just because I am in a good mood.*" (P4, C9). Sometimes, participants seek items not only for themselves but also for others. Three participants mentioned the target audience (three occurrences), e.g., P1 liked the explanation when it mentioned the kid, "*The recommended app could also be used by my kid to learn Chinese, ask questions, listen to songs, etc.*" (P1, C45). In addition, whether the user has a strong preference (25 occurrences) was mentioned by four participants. If a participant has no strong preference on the recommended item, they might follow the recommendation even though the explanation is simple, for example, "*As I have no idea of which one to buy, I get this suggestion from my kid that helps me make a decision.*" when the explanation is "*I like this cake because it is in yellow.*" (P1, C18). On the contrary, if a participant has a strong preference, the explanation may have less effect. For example, when P5 received an explanation about how tasty the rabbit was, P5 stated, "*I have never eaten rabbit, and I resist eating it, even though it is tasty.* (P5, C120).

Relationship between users and recommenders (Relation). Different from the relationship in the "relation-based" explanation, which contains the relationship between the user and the person(s) mentioned in the explanation, the "relation" here refers to the relationship between user and recommender. The relationship between user and recommender has built before receiving a new recommendation may affect the user's feeling on explanations. Four participants have a positive attitude toward the explanation when they have trusted the recommender (10 occurrences), for example, "*The product is recommended from the person I trust.*" (P5, C64). Moreover, all participants considered the common interest with the recommender (11 occurrences). They thought those good explanations were often from the person who has a common interest with them, for example, "*The people I have followed share common interests with me, e.g., same research topic or movie review. The contents that they liked interested me since they selected those contents.*"(P2, C124). In addition, all participants judged the explanation if the recommender has successfully/unsuccessfully recommended items to the user (nine occurrences), for example, "*My friend has taken me to eat some delicious food in HK so when she recommends me the restaurant or dessert shop, I will trust her taste.*" (P3, C10).

4.1.3 Recommender Attributes. We conclude some factors related to the recommender itself as the recommender attributes with their descriptions and examples (see Table 5). The "politeness" of the recommender made users feel comfortable. On the contrary, impolite behavior has the opposite effect. For instance, P1 had negative feelings when the salesperson was pushy, "*Let me feel not comfortable and too aggressive.*" (P1, C4). In addition, the recommender tries to consider the user context/preference while increasing the feeling of "benevolence". Comparably, two participants felt the recommender lacked "benevolence" when they received a "non-personalized" explanation. For example, P1 wrote, "*It seems she did not attempt to understand what I need, but just wants to earn money.*" (P1, C4). Two participants also doubted the "integrity" of the recommender when they received proofless "personal opinions". For example, P5 distrusted "*The conclusion 'It has less additive' from the second friend is just her speculation without any scientific evidence.*"(P5, C81). The personal opinion that is contrary to P2's view "*makes me disappointed*" (P2, C138). Besides, two participants showed that the "experience" of the recommender also increased the trustworthiness of the *personal opinions based explanation*. For example, P1 wrote, "*As the recommender has much experience of eating Japanese foods and is also familiar with this restaurant, his recommendations should be useful.*"(P1, C61). Moreover, two participants trusted the recommender's "expertise" if s/he had knowledge of the recommended item. For example, P2 felt good about the explanation because "*The article is written by one person in an NLP research group.*" (P2, C78). Finally, two participants were interested in the recommender who has a reputation. For instance, P5 wrote, "*I had heard about how successful the male beauty blogger Li Jiaqi is. He is called 'the man who sells most lipsticks in China'. So I was curious how he would recommend the product.*" (P5, C66).

4.2 Relationship Exploration

We provide an overview of the influence of user contexts and recommender attributes on social explanations in a network diagram (see Figure 1). Furthermore, Table 6 lists the relationship between social explanation categories and the detailed factors of user contexts and recommender attributes with the percentage of occurrences for each relationship.

4.2.1 Relation-Based. As an external context, "companion" (12.5%) reflects if the participant views recommendations with other people or alone, which can change a participant's attitude toward explanations, as the participant may consider the opinions of the companion, i.e., "*It was liked by my daughter*" (P1, C97).

Moreover, a participant felt the explanation matched the "preference" (12.5%), "*I like reading reviews and articles on XX*", which is a website mentioned in the explanation, "*The following are recommended by users of this website.*" (P2, C74). Furthermore, users'

Table 5: The classification of recommender attributes in the codebook with examples

Category	Description	Example
Politeness	The attitude to the user, such as polite, unaggressive, etc.	*"the salesperson's attitude is polite, patient, and not aggressive"* (P1, C1)
Benevolence	Whether the recommender (tries to) considers the user context/preference to provide a suitable item	*"The sales person did not recommend expensive goods to me, but tried to find sths that are really suitable for me."* (P1, C1)
Integrity	Whether the recommender tells the truth	*"the recommendation is sincere"* (P5, C128)
Experience (human)	The recommender has experience of the item or similar item(s)	*"My friend did make progress in her programming skills by using this platform."* (P4, C58)
Expertise	The recommender has knowledge of the item or similar item(s)	*"CloudMusic is the most successful music app that utilizes personalized recommendation algorithm to recommend songs or song lists to users."* (P2, C11)
Reputation	The recommender has received positive comments from other users	*"I had heard about how successful the male beauty blogger Li Jiaqi is."* (P5, C66)

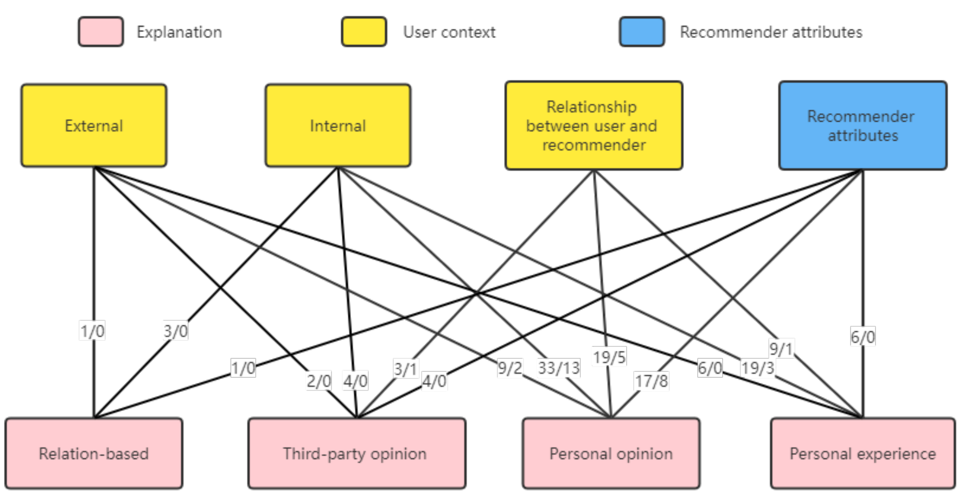

Figure 1: The map linking the social explanations with factors of user contexts and recommender attributes. The number in the lines represent the number of occurrences of good/bad explanation for the relationship.

"relevant experience" may change their attitude to the relation-based explanation (25%), e.g., "*I place more trust in the blogger who I have followed.*" (P5, C72).

In respect of recommender attributes, a participant mentioned "politeness" (12.5%) that made her feel good, "*The salesperson's attitude is nice and polite, so I felt comfortable.*" (P1, C97).

In addition, a participant was concerned about the closeness of the relationship between the person mentioned in the explanation and the participant, i.e., "*For some WeChat friends who are important to me, I may care what kind of article they would like to give the 'Wow' tag sometimes.*"(P5, C40). It indicates that the relationship between a user and the person mentioned in the explanation may be the key to influencing the user's attitude towards relation-based explanations.

4.2.2 Third-Party Opinions. The third-party opinions are linked to"location" of external situations (10%). "*It is a surprise recommendation for me since that place is in my hometown, but I have no idea about that place.*" (P3, C47). Although the third party has no relationship with the users, the relationship between users and recommenders can influence users' attitudes toward the third-party opinions. The participants believed in the explanation because of the established trust relationship with the recommender (15%), "*Although the comment is from an unknown person, I knew it from*

Table 6: The percentage of connections among social explanations and the factors that contribute to the users' feelings

Category	Relation-based	Third-party opinions	Personal opinions	Personal experiences
User Contexts				
External				
Environmental Condition	0	0	4.5%	2.3%
Companion	12.5%	0	1.1%	2.3%
Time	0	0	3.4%	0
Location	0	10%	3.4%	9.1%
Internal				
Demographics	0	0	4.5%	6.8%
Personality	0	0	4.5%	0
Physical	0	0	4.5%	6.8%
Domain Knowledge	0	0	1.1%	2.3%
Relevant Experience	25%	5%	10.1%	11.4%
Emotion	0	5%	1.1%	0
Target Audience	0	5%	1.1%	2.3%
Preference	12.5%	5%	24.7%	20.5%
Relation				
Trusted Relationship	0	15%	7.9%	6.8%
Common Interest	0	5%	10.1%	9.1%
Successful Experience	0	0	9.0%	6.8%
Recommender Attributes				
Politeness	12.5%	0	6.7%	0
Benevolence	0	5%	9.0%	2.3%
Integrity	0	0	2.2%	0
Experience	0	5%	6.7%	9.1%
Expertise	0	5%	2.2%	2.3%
Reputation	0	5%	1.1%	0

my friend, who I trust." (P4, C33). Besides, other influencing factors of user context (e.g., successful experience, emotion, target audience) and recommender attributes (e.g., benevolence, expertise, reputation) have a loose relationship with recommendations (the percentage is equal to 5%).

4.2.3 Personal Opinions. As the most common social explanation category in our participants' daily lives, "personal opinions" was influenced by all the factors we identified, especially those under internal context. The "preference" is the most frequently mentioned user context by participants (24.7%). When users have a strong preference, the recommender's personal opinions can provide information that helps users compare with their preferences. For example, P4 praised the recommender's personal opinion matching her preference, "*It is challenging [for the recommender], so I think the scenario [of the game] may be interesting.*" (P4, C55). When users have no strong preference, they are more likely to consider and accept the recommender's opinion. For instance, P1 wrote, "*As I have no idea of which one to buy, I got this suggestion from my kid that helps me make a decision.*" (P1, C18). Moreover, users are more likely to understand the recommender's opinions when they have relevant experience (10.1%). Three participants think the explanations related to their "relevant experiences" are good. For example, P3 understood the recommender's opinion because "*We have tried the cookie.*" (P3, C14).

In addition, users' attitudes on personal opinions can be influenced by their relationship with the recommender. Particularly, three participants have a positive feeling of personal opinion because they have a common interest with the recommender (10.1%). For example, P3 acknowledged the recommender's positive opinion because "*We have similar tastes in music and we always share music with each other.*" (P3, C53). Comparably, users tend to have a negative perception when the recommender has different interests or tastes from them. Three participants hesitated to make a decision because the opinions come from the recommenders who have other interests or uncertain preferences. For example, P1 stated, "*I am interested in watching this movie, but because the recommender does not have common interests with me, I am a little hesitant whether I will like it as her*"(P1, C62). Besides, other factors under user contexts and recommender attributes have a relatively weak relationship with the user's feelings on personal opinions (the percentage is less than 10%).

As shown in Figure 1, most of the "bad" explanations reported in the diary study are personal opinions due to neglecting or mismatching the user contexts or recommender attributes. Therefore, we need to consider these factors while providing personal opinions as explanations for recommendations.

4.2.4 Personal Experiences. The personal experiences emphasize the practical value of explanations based on human recommenders' experiences using the recommended item, which could increase the trustworthiness of the recommender. Similar to the findings of personal opinions, "preference" (20.5%) and "relevant experience" (11.4%) under the user contexts are the most influencing factors that can influence users' attitudes towards the explanations of personal experiences. For instance, the experience shared in the explanation can be helpful if it matches the user's preference, "*My friend's experience reminded me that [...] The spray plaster solves the problem.*"(P5, C130). Besides, participants appreciated the explanation that matched their relevant experience, "*She mentioned two impressive advertisements, which aroused my memory of fluffy after using shampoo.*" (P5, C111). However, the relationships between explanations and other factors such as environmental conditions, domain knowledge, and expertise are relatively weak (the percentage is less than 10%).

5 DISCUSSION

The results show that users' attitudes towards explanations may change due to a variety of factors. Therefore, a conversational recommender system can capture both external and internal user contexts and then adjust recommendations and explanations in time through continuous dialogue, which could help users make better decisions and build trust with users. Although most of the factors have been considered in context-aware recommender systems [1], we focus on their effects on user attitudes toward social explanations and try to explore the relationship between social explanation categories and the factors.

The majority of social explanations are related to personal opinions and experiences, which represent the own information provided by the recommenders. From the recommenders' perspective, they build rapport with users by sharing their personal opinions and experiences as a part of self-disclosure [2, 10]. However, our diary entries show that not all personal opinions and experiences can be good explanations. The users' attitudes toward personal opinions and experiences may depend on user contexts and recommender attributes. Therefore, **we suggest offering personal opinions and experiences in explanations tailored to user contexts, such as their relevant experiences, preferences, and common interests.**

Particularly, the effects of "preference" and "relevant experience" are evident in personal opinions and personal experience. These two factors are highly related because the previous positive experience is usually associated with a strong user preference. In most cases, users having strong preferences or relevant experiences are more likely to be convinced by social explanations, as users tend to perceive higher satisfaction when they are familiar with the recommendations [14]. However, sometimes the user may also have a negative experience with items similar to the recommendations, which may negatively influence the user's feeling about the explanations. Therefore, **it would be wise to provide personal opinions and experiences as explanations after it confirms that the user did not have a negative impression of the features of recommendations**.

Although the explanations based on third-party opinions are uncommon in our dairy study, we still foresee the value of such explanations, especially when the user trust the recommender. Like the social relationships among users, a trust relationship could facilitate the design of social recommender systems [5]. Moreover, the trust-inspiring explanation could also save users' cognitive effort and increase users' intention to use the recommender system in the future [23]. **Thus, we suggest trying to build a certain relationship between the user and the recommender, for example, common interests, before offering third-party opinions.**

A previous study shows that the recommendation explanations provided by users have a higher quality than the explanations generated by systems [17]. Therefore, in our dairy study, we aim to provide design implications for the explanations by recording and analyzing the explanations provided not only by systems but also by humans. In the result section, we also highlight some factor items specific to a human recommender. For example, the item "common interests" only applies to a human recommender since it does not make sense to explain as a system shares common interests with the user. Although the recommender system could learn how to provide richer explanations from the user-generated explanations, we still need to consider if the explanations match the system characteristics. Otherwise, the mismatched explanations may cause the "uncanny valley" effect [8] in the interaction between humans and recommender agents.

6 LIMITATIONS AND FUTURE WORK

There are two major limitations of this study. First, our study has a relatively small sample size according to the ideal sample size for a diary study (more than 10 participants) [13]. Nevertheless, our study results in substantial diary entries to explore social explanations. Second, since all the participants are knowledgeable about recommender systems, their views may differ from layman's on some social explanations. In the future, we plan to validate our research findings by evaluating the four categories of social explanations with varying contexts of user and recommender attributes in a conversational recommender system.

7 CONCLUSION

To explore social explanations in a broader view, we conducted a two monthly diary study of the explanations for recommendations that appeared in participants' daily lives, either from humans or systems. These recommendations covered multiple application domains, such as recommendations for products, music, and media. Furthermore, we evaluated four typical categories of social explanations regarding users' ratings and investigate how user contexts and recommender attributes may contribute to "good" or "bad" social explanations. Specifically, the explanations based on personal opinions and experiences are usually associated with internal user contexts (e.g., users' preferences and relevant experiences). Finally, based on these preliminary findings, we provide implications for designing personalized social explanations to enhance the rapport between users and recommender systems.

ACKNOWLEDGMENTS

The work was supported by Hong Kong Research Grants Council (RGC/HKBU12201620) and Hong Kong Baptist University IRCMS Project (IRCMS/19-20/D05). We also thank all participants for their time in taking part in the study and reviewers for their constructive comments on our paper.

REFERENCES

[1] Gediminas Adomavicius and Alexander Tuzhilin. 2011. Context-aware recommender systems. In *Recommender systems handbook*. Springer, 217–253.

[2] Irwin Altman. 1973. Reciprocity of Interpersonal Exchange. *Journal for the Theory of Social Behaviour* 3, 2 (1973), 249–261. https://doi.org/10.1111/j.1468-5914.1973.tb00325.x

[3] Svetlin Bostandjiev, John O'Donovan, and Tobias Höllerer. 2012. TasteWeights: a visual interactive hybrid recommender system. In *Proceedings of the sixth ACM conference on Recommender systems*. 35–42.

[4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[5] Chaochao Chen, Jing Zeng, Xiaolin Zheng, and Deren Chen. 2013. Recommender system based on social trust relationships. In *2013 IEEE 10th International Conference on e-Business Engineering*. IEEE, 32–37.

[6] Li Chen and Luole Qi. 2010. A Diary Study of Understanding Contextual Information Needs during Leisure Traveling. In *Proceedings of the Third Symposium on Information Interaction in Context* (New Brunswick, New Jersey, USA) *(IIiX '10)*. Association for Computing Machinery, New York, NY, USA, 265–270. https://doi.org/10.1145/1840784.1840823

[7] Robert B Cialdini and Lloyd James. 2009. *Influence: Science and practice*. Vol. 4. Pearson education Boston.

[8] Leon Ciechanowski, Aleksandra Przegalinska, Mikolaj Magnuski, and Peter Gloor. 2019. In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems* 92 (2019), 539–548.

[9] Georg Groh, Stefan Birnkammerer, and Valeria Köllhofer. 2012. Social recommender systems. In *Recommender systems for the social web*. Springer, 3–42.

[10] Shirley Anugrah Hayati, Dongyeop Kang, Qingxiaoyang Zhu, Weiyan Shi, and Zhou Yu. 2020. INSPIRED: Toward Sociable Recommendation Dialog Systems. *CoRR* abs/2009.14306 (2020). arXiv:2009.14306 https://arxiv.org/abs/2009.14306

[11] Jianming He and Wesley W Chu. 2010. A social network-based recommender system (SNRS). In *Data mining for social network data*. Springer, 47–74.

[12] Chin-Chi Hsu, Mi-Yen Yeh, et al. 2018. A general framework for implicit and explicit social recommendation. *IEEE Transactions on Knowledge and Data Engineering* 30, 12 (2018), 2228–2241.

[13] User Interviews Inc. 2022. The User Experience Research Field Guide. https://www.userinterviews.com/ux-research-field-guide-chapter/diary-studies Accessed = 2022-04-22.

[14] Dietmar Jannach, Lukas Lerche, and Michael Jugovac. 2015. Item Familiarity as a Possible Confounding Factor in User-Centric Recommender Systems Evaluation. *i-com* 14, 1 (2015), 29–39. https://doi.org/doi:10.1515/icom-2015-0018

[15] 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 (2012), 441–504.

[16] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2019. Personalized explanations for hybrid recommender systems. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. ACM, New York, NY, USA, 379–390. https://doi.org/10.1145/3301275.3302306

[17] Johannes Kunkel, Tim Donkers, Lisa Michael, Catalin-Mihai Barbu, and Jürgen Ziegler. 2019. Let Me Explain: Impact of Personal and Impersonal Explanations on Trust in Recommender Systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300717

[18] Mary L McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia medica* 22, 3 (2012), 276–282.

[19] Martijn Millecamp, Nyi Nyi Htun, Cristina Conati, and Katrien Verbert. 2019. To Explain or Not to Explain: The Effects of Personal Characteristics When Explaining Music Recommendations. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 397–407. https://doi.org/10.1145/3301275.3302313

[20] Cliodhna O'Connor and Helene Joffe. 2020. Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods* 19 (2020), 1609406919899220. https://doi.org/10.1177/1609406919899220 arXiv:https://doi.org/10.1177/1609406919899220

[21] Alexis Papadimitriou, Panagiotis Symeonidis, and Yannis Manolopoulos. 2012. A generalized taxonomy of explanations styles for traditional and social recommender systems. *Data Mining and Knowledge Discovery* 24, 3 (2012), 555–583.

[22] Florian Pecune, Shruti Murali, Vivian Tsai, Yoichi Matsuyama, and Justine Cassell. 2019. A Model of Social Explanations for a Conversational Movie Recommendation System. In *Proceedings of the 7th International Conference on Human-Agent Interaction* (Kyoto, Japan) *(HAI '19)*. Association for Computing Machinery, New York, NY, USA, 135–143. https://doi.org/10.1145/3349537.3351899

[23] Pearl Pu and Li Chen. 2007. Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems* 20, 6 (2007), 542–556.

[24] John Rieman. 1993. The Diary Study: A Workplace-Oriented Research Tool to Guide Laboratory Efforts. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems* (Amsterdam, The Netherlands) *(CHI '93)*. Association for Computing Machinery, New York, NY, USA, 321–326. https://doi.org/10.1145/169059.169255

[25] Kate Roberts, Anthony Dowell, and Jing Bao Nie. 2019. Attempting rigour and replicability in thematic analysis of qualitative research data; A case study of codebook development. *BMC Medical Research Methodology* 19, 1 (2019), 1–8. https://doi.org/10.1186/s12874-019-0707-y

[26] James Schaffer, John O'Donovan, James Michaelis, Adrienne Raglin, and Tobias Höllerer. 2019. I Can Do Better than Your AI: Expertise and Explanations. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 240–251. https://doi.org/10.1145/3301275.3302308

[27] Young-Duk Seo, Young-Gab Kim, Euijong Lee, and Doo-Kwon Baik. 2017. Personalized recommender system based on friendship strength in social network services. *Expert Systems with Applications* 69 (2017), 135–148.

[28] Amit Sharma and Dan Cosley. 2013. Do Social Explanations Work? Studying and Modeling the Effects of Social Explanations in Recommender Systems. In *Proceedings of the 22nd International Conference on World Wide Web* (Rio de Janeiro, Brazil) *(WWW '13)*. Association for Computing Machinery, New York, NY, USA, 1133–1144. https://doi.org/10.1145/2488388.2488487

[29] Anjeli Singh and Sareeka Malhotra. 2013. A Researcher's Guide to Running Diary Studies. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction* (Bangalore, India) *(APCHI '13)*. Association for Computing Machinery, New York, NY, USA, 296–300. https://doi.org/10.1145/2525194.2525261

[30] Timothy Sohn, Kevin A. Li, William G. Griswold, and James D. Hollan. 2008. A Diary Study of Mobile Information Needs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 433–442. https://doi.org/10.1145/1357054.1357125

[31] Zhoubao Sun, Lixin Han, Wenliang Huang, Xueting Wang, Xiaoqin Zeng, Min Wang, and Hong Yan. 2015. Recommender systems based on social networks. *Journal of Systems and Software* 99 (2015), 109–119.

[32] Maxwell Szymanski, Martijn Millecamp, and Katrien Verbert. 2021. Visual, Textual or Hybrid: The Effect of User Expertise on Different Explanations. In *26th International Conference on Intelligent User Interfaces* (College Station, TX, USA) *(IUI '21)*. Association for Computing Machinery, New York, NY, USA, 109–119. https://doi.org/10.1145/3397481.3450662

[33] Chun-Hua Tsai and Peter Brusilovsky. 2019. Explaining Recommendations in an Interactive Hybrid Social Recommender. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 391–396. https://doi.org/10.1145/3301275.3302318

[34] Xiwang Yang, Yang Guo, Yong Liu, and Harald Steck. 2014. A survey of collaborative filtering based social recommender systems. *Computer communications* 41 (2014), 1–10.

[35] Yuhang Zhao, Shaomei Wu, Lindsay Reynolds, and Shiri Azenkot. 2018. A Face Recognition Application for People with Visual Impairments: Understanding Use Beyond the Lab. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173789