Evaluating Explanations for different Relationship Strengths

Master's Thesis

Nivedita Prasad

Evaluating Explanations for different Relationship Strengths

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

COMPUTER SCIENCE
TRACK DATA SCIENCE

by

Nivedita Prasad

4712099

N.Prasad@student.tudelft.nl

Web Information Systems
Department of Software Technology
Faculty EEMCS, Delft University of Technology
Delft, the Netherlands
http://wis.ewi.tudelft.nl

Abstract

People like to travel in groups to visit places. Group recommendation systems can be used to recommend an itinerary of "places of interests" (POIs) in an ordered sequence. The order of POIs in the sequence can be explained to group members to increase acceptance of the recommended items. There is a possibility that explanations which reveal names and rating preferences could create a threat to privacy.

The main study in this work uses two group types - a primary group consisting of closely-related members, and a secondary group consisting of loosely-related members. Explanations with either complete information or privacy-preserving information are offered alternatively to these groups. The purpose of this study is to evaluate whether different group types need different types of explanations to improve their satisfaction. These explanations explain the entire recommended sequence of POIs with regard to possible conflicting situations that could occur due to disagreement with the order of the sequence.

A total of 25 participants took part in the evaluation. There was no significant difference identified between the explanation types preferred by each group type. To understand the underlying reason for this result, a post-hoc analysis was done. We identified a participant's most frequently used conflict-handling modes using the Thomas-Kilmann personality assessment test. We then analysed the user comments provided during the questionnaire. The analysis potentially suggests that different conflict-handling modes could be a factor affecting which explanation type was preferred by a person when they are in a particular group (e.g. primary vs secondary).

Thesis Committee:

Chair:	Prof. Dr. Ir. A. Bozzon, Faculty EEMCS, TU Delft
University supervisor:	Prof. Dr. Nava Tintarev, Faculty EEMCS, TU Delft
Daily supervisor:	Ir. Shabnam Najafian (PhD Student), Faculty EEMCS, TU Delft
Committee Member:	Prof. Dr. Odette Scharenborg, Faculty EEMCS, TU Delft

Acknowledgement

My two years in TU Delft have taught me more than what life could ever teach me. I have learned to push my boundaries further, realize my caliber, and work harder. My experiences at the TU have molded me to be a better person.

A thesis is never the work of a single person. I would like to thank my thesis supervisor Dr. Nava Tintarev for guiding me constantly. Her eye for detail and her clear approach to addressing problems have helped me a lot in structuring my thesis.

I would like to thank my daily supervisor, Ir. Shabnam Najafian (Ph.D. student), who has guided me through the entire journey of my thesis. She has mentored me with much feedback and suggestions, and has made sure that I never lost the originality of my work.

I would like to thank the members of the Epsilon group and the participants who took part in my survey and willingly gave their precious advice during the interview process. The user studies conducted would not have been possible without their cooperation.

I express my profound gratitude to Arun Gunda for truly being my better half during this roller coaster ride and without whom this thesis would not have been possible. Also, I would like to thank Naveen Srivatsav and Niharikha Mahajan for being my extended family at Delft and constantly providing me support throughout the two years of my study here in TU Delft. I would like to thank my good friends Aishwarya Shastry, Enreina Annisa Riziasri and Sandesh Manganahalli Jayaprakash to have guided me during my thesis. Finally, I would like to thank my parents, the rest of the family and of course my dogs without whom this dream of mine would have never come true. Thank you for believing in me.

Nivedita Prasad
Delft, the Netherlands
August 24, 2019

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

In this thesis, we want to evaluate if different types of explanations given to groups with different relationship strengths affect the satisfaction of members of the group concerning the recommended items. We want to explore if a set of closely related people prefer *transparency* or opt to preserve their identity when group's preferences are explained. Similarly, we want to also investigate if a group of people who do not share a personal bond would opt for *privacy-preserving* explanations to increase their satisfaction in comparison to transparent explanations.

Recommendation Systems (RS) recommend items to an individual or group of individuals but their underlying working is generally unclear to the normal user. These recommendations can be explained [44]. "Explanations are justifications or descriptions to make something clear to a person" [40]. The RS attempts to give an explanation to an individual or a group to explain the various underlying reasons for the selection of the recommended item(s). These reasons could be the constraints undertaken to select the item [33], the similarity of item's content [6], or similarity of users who have also opted for the same recommended item when they had the same need [12]. The explanations can also be based on the underlying aggregation algorithm, by explaining the way predictions were aggregated for recommending items for a group [35, 27] or explain recommendations based on the social interactions within the group [38]. These explanations are given in terms of text or visualizations to the user [12].

It is relatively simpler for a RS to make decisions for a single user compared to recommendations for a group. An example would be helping a single user select a book. When a group of people want to travel to a sequence of places or dine together in a restaurant, a group decision should be taken. "A decision problem arises when group members have varied opinions about the correct choices to make" [13]. When groups are recommended a list of places, songs or movies, the group has to be kept satisfied at all times. It is difficult to arrive at a consensus or be fair to every group member [35]. This issue can be mitigated by explaining how the preferences of individual group members are taken into account to make the group aware of their group member's constraints [35, 27, 12].

Explanations given to the individual user can be classified into one of the seven aims formulated in [45]. These aims are *transparency, satisfaction, scrutability, trust,* *effectiveness, persuasiveness, efficiency*. In group recommendations, the individual predictions or models are aggregated to make a recommendation to a group of people who want to perform the same activity [31]. Then the explanation given should not only include the seven aims discussed above, but also one or more of the group aspects like a consensus, optimality, fairness. For example, some state-of-the-art group travel-recommendation systems can generate a sequence of POI (places of interest) for groups to visit [22, 27, 3]. The group has to be kept satisfied throughout the sequence recommended to them by helping them reach consensus through explanations. Explanations can also involve group dynamics (relationship types, conflict style).

Current work on group recommendations is focusing on improving recommendations by incorporating the group's attributes (social factors) [31]. These attributes are the group's relationship strength (strong ties like couples or close friends, weak ties like acquaintances) [16], personality of each group member [37], type of relationship (groups where you "share everything", "respect other member highly", "equal footing" with other group members, "compete with other group members") [31], expertise in a particular domain [16], and demography (children, adults with disability, adults) [3]. Explanations have also incorporated these social factors (social explanations) to help group members understand the social reality of the group and improve the acceptance of the recommended items [38].

Explanations with the aim of consensus can be designed for closely related groups like friends or couples without involving the above-mentioned social components. These explanations have increased the satisfaction of the group members [35, 27]. Note that these explanations are transparent with the names and ratings of group members when they report the member who was responsible for the lowest and highest ratings (if that was a reason for a recommendation). There is no work on whether transparent explanations will be accepted by group members who are acquaintances. They may insist on hiding their identity (preserve privacy) from other group members to maximize their satisfaction with the given explanation [27]. It is important to investigate whether group members need this privacy preserved type of explanation.

The literature until now has focused on generating explanations that might or might not involve social components for an *individual item* in the sequence. There has been no work that has designed explanation for the *entire* sequence that could be presented to the user. Also, there is less literature on how all the items in the sequence should be explained and the order in which these items should be explained. Until now, the aggregated preference [35, 27] or the aggregated model [12] has been explained to the group members. This is done by explaining (1) the algorithm used (2) information about the names of members who were responsible for the highest or lowest rating of an item, (3) information on group's social reality, (4) information about similar users or similar content. These points are used in combination or individually in an explanation.

It is important to conduct preliminary studies to determine the content of the explanation and how this information can be presented to the individual user for the whole sequence of POIs. The utility of explaining the sequence is that it gives a holistic outlook of the entire trip that the recommendation system has suggested to the user. Furthermore, it has been suggested that, in the domain of tourism, explaining the whole sequence of places (complete itinerary) would be preferred by tourists to improve their user satisfaction versus explaining a single item from the sequence [27].

The following sections mentioned in this chapter contains some technicalities that would be well understood as the reader reads into the chapters explained.

1.2 Research Goal

The research questions addressed in this thesis are listed below.

1. **RQ 1: How should one formulate explanations for user satisfaction?**
This question investigates the content to be presented in the explanation. Explanations that contain information about aggregated item preferences along with names of group members and item names are investigated. These explanations are designed for situations that may cause conflict when the user compares their own preferences with the order of the recommended sequence. Possible situations that may cause conflict are considered in this study. Explanations are designed for each situation.

2. **RQ 2: How should one formulate and structure an explanation for different relationship strengths?**
By using the explanations designed in the previous question for a particular situation, this question investigates the structure of the explanation for the entire recommended sequence with the support of user comments acquired via a user study. It also investigates if there a need to change (i.e., preserve privacy) the formulated content of the explanation according to the relationship strength.

3. **RQ 3: Do different explanation types for different relationship strengths influence user satisfaction?**
The presence or absence of names is used to measure the satisfaction value of the explanation given for the recommended sequence. Additionally, these values are supported by relevant user comments from the participants. This question investigates how the two explanation types (transparent and privacy-preserving) given for the entire sequence (four possible social situations combined) affected an individual user's satisfaction when they belonged to groups with different relationship strengths.

1.3 Results

Each research question was answered by conducting a user-specific study.

1. **RQ 1:** A user study was conducted where the user comments were analyzed for determining the content for the explanation. The content was determined in terms of word choice, the effect created by the explanation, usage of categorical values in place of real numbers, and conveying facts that were pleasing and displeasing to achieve better user satisfaction.

2. **RQ 2:** A user study was conducted where the inferences on the content of the explanation, potential structure, and the privacy required by the users are made

from user comments. The disclosure of names made the explanations look personalized or, in contrast, created social discomfort amongst friends. With acquaintances, disclosure of names seems to create a breach of privacy. The inference on the structure was that the information about favorite places being visited could be explained first. Next, information about system transparency could be given after algorithmic transparency. Finally, assuring information about compromises made by other group members gave a holistic approach to the explanation.

3. **RQ 3:** An evaluation study was conducted.

   - We found that the content and structure of both the transparent and privacy-preserving explanation types have an overall satisfaction that is above 3.6 on a Likert scale of 5. Hence the explanations we have designed have satisfied the participants well.
   - There was no significant difference between the explanation types a group type preferred. Hence we could not prove that the primary group (close friends) preferred transparent explanations and secondary group (acquaintances) preferred privacy-preserving explanations.
   - By conducting posthoc analysis, we intend to find the potential reason behind high standard deviation found across all the four scenarios. A Thomas-Kilmann instrument personality test for determining the personality profile was conducted. The user comments segregated under each conflict mode seemed to follow a pattern.

1.4 Contribution

The contribution of this work is listed below.

- Designing explanations for the entire recommended sequence. The explanations contained information in terms of situations where the individual can(not) have a conflict with the recommended order of the sequence. This information is arranged in a "sandwich model" [11] to increase satisfaction.

- An explanation containing the information about the names and rating preferences of the group members is called a transparent explanation. The names of the members are hidden to make it a privacy-preserving explanation. We have evaluated whether the satisfaction of users belonging to different relationship strengths was influenced differently after receiving these two different explanation types.

1.5 Thesis Outline

A survey of the existing literature has been presented in Chapter 2. It contains a survey of different explanation types, aggregation strategies, and group attributes. In the main work, three experiments have been conducted. The first two experiments are "preliminary studies" that were conducted before the final evaluation. They are explained in

Chapter 3. Preliminary study 1 discusses how the supporting user comments helped us determine the content to be presented in an explanation for a group member. Preliminary study 2 discusses how the supporting user comments help us order and plan the information we have obtained into a final structure. This structure can accommodate the explanation of the entire recommended sequence. The final explanation designed for the entire sequence is given to participants in a within-subject study. The participants will validate the given explanation across four scenarios explained in Chapter 4. The conclusions and future work are found in Chapter 5.

Chapter 2

Literature Survey

The research question we are answering in this thesis is whether different explanation types given to different relationship strength influence user satisfaction. The group type is based on the relationship strength present amongst the group members. The group types are: the primary group is a closely related set of people and the secondary group does not share a personal bond. They can be a couple, close friends (primary group) or acquaintances (secondary group). The explanation designed is explaining the entire sequence of recommendation to every group member. Thus this work is built in two sections. They are (1) Designing explanations for a sequence generated by an aggregation algorithm and (2) Evaluating whether the designed explanations are influencing the user satisfaction.

To this end, we are looking in-depth about the related work necessary to solve the question. Firstly, we will look briefly on how aggregation strategies generate a sequence of recommendations for the group. Secondly, we will discuss the designing of explanations for groups and how the recommended items in the sequence are explained to groups. As a next step, we will discuss on group attributes particularly relationship strength and their influence in group RS. Furthermore, we will discuss on explaining to groups while considering these group attributes. In the end, we will look at the evaluation of explanations.

2.1 Strategies for generating sequences

The preferences that are procured from individual users are aggregated to make recommendations for a group [35, 24, 30, 26, 29, 31, 13]. Group recommendation algorithms can either (a) aggregate the items' rating preferences/items themselves for each individual user or (b) aggregate individual user-profiles and then make a group profile to recommend items for the group [15]. The former is called as *aggregated predictions* and the latter is called as *aggregated models*.

In this section, we are going to discuss on aggregation strategies and how ranked sequences are generated from them. The two basic types of aggregation strategies (aggregated models and aggregated predictions) are explained first in Section 2.1.1.1. Secondly, we have discussed the importance of ordering of a generated sequence of recommended items. Finally, the different types of aggregation strategies (Social choice strategies) are explored in Section 2.1.1.3. This is important because, in this work, we will be designing explanations without social components for a sequence that is based on an aggregation strategy and then evaluating whether they are influencing user satisfaction of group members having different relationship strength. Therefore it is vital to understand the different ways in which sequence can be generated using algorithms that are generally used to make decisions for a group (social choice strategies).

A group of people can get relevant recommendations when group recommendation approaches are applied to the individual group member's preferences that are given for the set of candidate items. The group recommendation approaches can be classified on the basis of their **characteristics** [15]. They are classified as follows: strategies of preference aggregation, recommendation algorithms, knowing user's preferences from previous recommendation-iterations, not knowing the preferences beforehand, then depending on whether the output of the group recommendation is a single item or a sequence, etc., and finally it can be based on how the preferences are collected - implicitly or explicitly.

Although there is a list of different group recommendation approaches mentioned above, we have focused only on the "strategies of preference aggregation" approach which is relevant for my thesis. In the following sections, we have discussed the types of aggregation strategies and different social strategies that generate sequences of recommended items for a group. These social strategies can be applied in the field of travel, music, movies, etc.

2.1.1 Aggregation Strategies

2.1.1.1 Types

In *aggregated predictions* strategy, user's preferences can be aggregated in two ways [15]. They are: (a) Either the recommended items are merged to form a set of solutions with *no ranking* from various solutions given, and from this the group selects a final list of recommendation [15]; or (b) individual member's rating for every individual item is aggregated and then *ranked* to form a sequence of recommended items as final result [15]. Therefore individual user's recommendations or raw ratings are aggregated and then a recommendation is given to the group in the form of sequences that are ranked or a list that is not ranked.

Aggregated models give the members of the group the opportunity to reduce the *privacy concerns* of the group [15]. This is done by creating a group profile and the recommendations are given to the group profile. Therefore the user profiles are aggregated into a group profile as said before and then recommendations are given to that group profile.

In the next section, we will discuss in brief about the importance of ordering of items and the social strategies for aggregated predictions. This is because we are going to use a ranked recommended sequence as a base to generate explanations for groups having two different relationship strengths.

2.1.1.2 Order of recommended Items

It is important to know *"why"* to rank/order the outcome of the sequence before discussing social strategies applied to aggregate predictions.

Masthoff in [31] has explained that the order of the sequence is important when a sequence of songs are played from one genre and there is a need to shift to another genre. Here there is a need to maintain the rhythm of the songs played and the mood of the users listening to it to increase satisfaction [31]. In the domain of news, the reader might not want to read sad news in-between two happy news items, hence there is a need to follow a specific order in the sequence of presented items.

This is suggested because the "overall satisfaction" of the given sequence is affected by the order of items [31]. Therefore, three reasons that are important for the order of items in the sequence are : (1) Consistency of the mood of the users affected by the item presented next (Example: list of songs), (2) Items at the end of the sequence should be a item that is worth being remembered by the users (Example: list of songs), (3) Item that is related to each other to be presented together (Example: news about sports presented together).

The author in [31] has also suggested that the outcome of an aggregation strategy is a sequence that can be ordered in descending, ascending or random order. It can be a list of POI for making an itinerary or a list of songs for a playlist based on the combination of preferences of individual group members. This order of the items in the sequence is important to the overall satisfaction of the group [35, 27].

These items in the sequence can be explained [12]. For example, for a sequence of songs explaining why a favorite song (rated high) of a group member is not been selected yet is important to be explained to keep that group member satisfied [35]. The other reasons when the explanations are given were when a favorite song was never selected or when a song they dislike is being played. It is also possible to face these kinds of situations where there is dissimilarity between the recommended sequence and the individual's own interest. A similar situation of disagreement can be faced when an itinerary is being recommended to a group [27]. All these works explain an item that is responsible for similar situations one by one when such a situation arises but never all situations at once. In the domain of tourism, literature has acknowledged that explaining the whole itinerary would be preferred by tourists to improve their user satisfaction versus explaining a single item from the sequence [27].

There is limited work that discusses the order in which the items in the sequence should be explained. This means there is not much literature that explains which item should be explained first or in a sequence even though the importance of sequence order is acknowledged. There is limited work in how an item in the sequence is explained to group members based on a social situation. Yet there is no work in explaining all the social situations that cause conflict and may arise due to a sequence recommended in a particular order.

The importance of the order in which the sequence of places of interests is explained given the recommended sequence has been considered in this thesis. In Section 3.3.3 of the Chapter 3, discusses the explanations that can be designed for different conflict situations. In brief, the group members may face conflicting situations where their preferences do not match the preference of the other group members. These situations are ordered such that it increases user satisfaction while the whole sequence is being explained.

2.1.1.3 Social Choice Strategies

The authors of [4] have stated that there is no optimal solution while aggregating individual preferences. In order to accommodate the preferences of individuals for the whole group, there are many social aggregation strategies [29]. In [15], the authors have categorized the aggregation functions into three groups. They are namely: "Majority based", "Consensus-based", "Borderline-based" aggregation functions.

The ***majority based*** functions considers items that are "very popular" [29, 15]. The social aggregation strategies that fall in this category are Plurality Voting, Borda Count and Copland rule [15]. The sequences are generated when the various definitions of the "majority" group mentioned here are applied to the user's ratings. Plurality voting is about choosing the item that gets the highest number of votes from the group members. Borda Count is where "Points are awarded to each alternative according to its position in the individual's preference list" [29]. Finally, the Copland rule orders alternative choices according to the "Copland index". The Copland index is defined as "the number of times an alternative beats other alternatives minus the number of times it loses to other alternatives" [29].

Consensus means an "agreement" that the whole group has to arrive by considering all the group member's preferences [29]. The social aggregation strategies that follow the definition of consensus are Additive Utilitarian, Average, Multiplicative, Average without Misery and Fairness. Using these strategies we can either select the highest outcome value as the final item or make a sequence by arranging the values from maximum to minimum or in random order. In Additive Utilitarian the sequence is arranged in the descending order *sum* of ratings of the candidate items. While in Multiplicative strategy, the sequences are arranged in the descending order of *product* of the individual user ratings for each of the candidate item. In the average strategy, the sequences are arranged in the descending order of maximum *average* to a minimum average of rating values calculated for one item over all the group members. Average without misery means the specific item with absolute rating values that are below a *threshold* are ignored from the sequence and then arranged in descending order of the average of rating values calculated for one item over all the group member. Finally, Fairness lets the individuals *take turns* to choose an item for the final sequence [35, 29, 13].

Borderline strategies generate sequence by considering only a subset of user preferences [15]. The aggregation functions which follow the definition of borderline strategies are Least Misery, Most Pleasure, Majority Voting, Most respected Person. In the least misery strategy, it makes sure "no-one is unhappy [15, 35]. Here the *maximum of the lowest rating value* is selected for every item and then the outcome is ordered from maximum to minimum. In most pleasure strategy, the *maximum of the highest rating value* is selected for each item and similarly, the outcome is ordered. Next, in majority voting, an item with the *highest number of votes* is the first in the sequence. The final strategy of the most respected person recommends the rating that has been suggested by the most respected of the group first and then the sequence is created.

It is possible to combine the above-said strategies to make a hybrid in order to counteract the disadvantages posed by each of them. In the work of [35] they have chained most pleasure, least misery and average without misery to make a robust aggregation algorithm that can suggest recommendations that can make no-one in the group unhappy.

The above-discussed aggregation predictions and aggregated models can be applied to the traditional algorithms of collaborative filtering, content-based, critiquing-based and constraint-based recommendations [15] to generate group recommendations. As a first step, the ratings and items are derived using the traditional algorithms and then aggregation strategies can be applied to them to produce a recommendation sequence for the group.

In the domain of tourism, these aggregation strategies are applied in order to generate a sequence of POI (places of interest) to the group [14, 27, 22]. In the work of [10], they observed that the participants mostly opted average strategy while having a discussion to decide the final places. In the work of [22], they have particularly applied the average, average without misery and most pleasure strategies to aggregate user's preferences. These strategies can also be applied to generate a sequence of songs to be consumed by a group [35].

The literature on explanations focus on understanding how individual members may relate to the group recommendations given. The sequence recommended to the group is sometimes not easily accepted by the group. They would expect reasons for situations where their own preference is not considered. This may arise due to the nature in which the social aggregation strategy has aggregated individual user's preferences. Hence explanations can be given to make each group member understand about other member's preference by explaining about the conflicting situations that may arise due to sequence recommended [35, 27].

2.1.2 Summary

The above sections at first discuss the definition and types of aggregation strategies. Secondly, we have seen the importance of the presentation of the sequence in a given order. Finally, we have discussed the working of each of the social aggregation strategies and how they generate sequences.

There is acknowledgment in the literature to explain the same to groups and individuals. The presence of explanations has increased the acceptance of the recommendations better [44, 37]. The following section will help us understand the different explanations that can be designed for individuals and groups.

2.2 Explanations

"Explanations are justifications or descriptions to make something clear to a person" [40]. They can be considered as additional information given to a user by the designer of the RS in addition to the recommendations given [12]. Another way to look at explanations is from the perspective of the user of the RS [12]. These explanations can be delivered in the form of text, visuals or both depending on their efficiency to deliver the aim of the explanation. Explanations designed with an aim will have criteria that they are trying to improve for a RS or the user. For example, transparent personalized social explanations are given to individuals for improving the persuasiveness, effectiveness, efficiency of Happy-Movie application and increase individual and group satisfaction [38].

It is necessary we delve more into the aim of transparency in this section. As we aim to increase user's satisfaction regarding the recommended items using an explanation and therefore reduce conflict among group members by explaining how the sequences were generated by the RS in this thesis

2.2.1 Goals and Guidelines

Explanations can have different goals that they want to achieve when they are presented to a user. The first step is to decide on **a goal** while designing an explanation. Some examples where explanations with goals are used are also mentioned in this section. The second step is the design guidelines [44] that an explanation designer has to follow.

Goals: The author of [45] from her research has identified and summarised seven different **goals**. One or more of these goals can be selected and explanations can achieve them when they are given to the user who has been recommended items. She has also identified in her work [45] that these explanatory aims could be "complementary or mutually exclusive (trade-off) to each other". Selecting an criteria /aim is the *first step* when a designer wants to design an explanation.

There are seven different explanatory aims as presented in the work of [45]. They are Persuasiveness ("Convince users to try or buy"), Satisfaction ("Increase the ease of usability or enjoyment"), Scrutability ("Allow users to tell the system it is wrong"), Effectiveness ("Help users make good decisions") [44], Efficiency ("Help users make decisions faster"), Trust ("Increase users confidence in the system") and Transparency ("Explain how the system works") [45]. It is not a necessity to select all of the seven goals. It is *not possible* for an explanation to achieve all the seven goals at once [45].

Explanations have these *goals* they want to fulfill when delivered to the users of the RS [45]. For example, users may need additional information from the RS to understand the recommendations better [12], understand why such a recommendation is given to them [45, 42, 7, 35] so that the system does not act as a black box. The explanations which explain "why-questions" are known as transparent explanations [45]. Herlocker et al., [20], has compared persuasive explanations which are tables of ratings, histograms, display of confidence level to find the most persuasive method to make a user choose a movie to watch. Explanations can be designed for increasing the satisfaction of the user by displaying the reason behind how the recommendation was given to users than being just being persuasive [5]. The work of [44] mainly concentrated on the aim of Effectiveness (help a user make a good decision on the item recommended) and compared the effectiveness of explanations versus satisfaction and persuasion.

Transparency is the clarification given to the user about how a recommender system has selected an item for a user(s) so that the recommender system is not perceived as a black box. The importance of transparency is experimentally validated in the work of [42] where the author tested five music RS by providing a transparent explanation versus non-transparent explanation to the participants for the songs recommended. The results supported transparent explanations as the users did not want blind recommendations even when they liked the song but not just requiring an explanation only when a new song is introduced to them. In the field of art, the authors of [7] have demonstrated the importance of transparent decision making process by displaying the percentage of sureness of the right decision made by the system and a "**why**" word, which shows a pop-up window listing the common properties of the recommended artworks the users' had positively rated. This is how they explain transparency visually to users to test the effects of the transparency on trust and acceptance of the content-based art recommender system. These are some examples where the goal of transparency was given importance.

Selecting goals: It is *hard* to create an explanation that does well on all the seven goals [45]. It is usually a "trade-off". If the explanation is long due to the criterion of transparency, it may not help in efficiency [45]. The examples cited are that the "system's goal" of trust is important to a book RS while the goal of user satisfaction is important for a television program recommender system. Similarly user satisfaction with the recommended sequence of an itinerary, is important for tourist. They may not bother about the efficiency of the system in selecting an itinerary. Therefore providing transparent explanation about how the system works could help in improving user's satisfaction with the recommended sequence.

Design Guidelines: In this paragraph, we are going to discuss some important guidelines that are necessary to be followed when an explanation is designed after a suitable explanation goal is selected. The author [44] has advised steps/guidelines to **design** the Explanations. The steps are as follows: (1) Explanations should have a *aim* they would like to achieve when they are presented to the user, (2) These explanatory aims should be evaluated by a *metric*, (3) The explanations generated for the items are affected by the metric used to evaluate the underlying recommendation *algorithm*, (4) The way we *present* recommended items to the user affects the explanation, (5) The model was chosen for *interaction* between the user and the recommended item affects the explanation presented, (6) The recommendation algorithm chosen affects the explanation we generate. We have fitted these guidelines to design our explanation which is seen in the next paragraph.

In this thesis, we are presenting transparent explanation to the group member and then measuring user satisfaction regarding the recommended items. These explanations will be designed for the entire sequence of places of interest with transparency (how the system works) as the aim. The underlying recommendation algorithm is chosen from the work of [35], where the ratings are aggregated using three social choice strategies that are chained together. Hence effectively the maximum and the minimum rating are averaged for each candidate item after the removal of items who have received ratings below a certain threshold (average without misery) to generate the final sequence. We present the explanation to the users and they evaluate them for user satisfaction via a questionnaire.

Factors affecting Explanations There can be factors that affect the explanations that RS give to the users. Tintarev in [44, 40], has explained that there are three factors *affecting* explanations namely: *presentation* of the items being recommended, the underlying *recommendation algorithm* predicting the items and *interaction* of the user with the items displayed by the system. In our thesis, the explanation could be affected by the underlying recommendation algorithm as we are trying to explain the sequences using transparent explanations. This is because the way in which predictions are aggregated are explained to the user.

In the next section, we will discuss very briefly about explanations for single user RS before moving for explanations given to groups (see Section 2.2.3).

2.2.2 Explanations for individual Users

In Section 2.2.1, we have discussed seven goals of explanations which are suitable for explanations given to individual users and the examples where they were used. In this section, we will look at sample explanations made for individual recommender systems and the domain they are used in. This is done to get a better idea about group explanations. We are particularly going to delve deep into textual explanations.

Explanations can be categorised based on the goals of explanations [12, 45] (see 2.2.1). These explanations can contain the basic information, additional information or explanation for "no solution found" scenario [12]. This means if the underlying recommendation algorithm is collaborative filtering for example then, the corresponding sample of basic explanation may look like "*Customers who bought an item a also bought item b*".

The explanations can give additional information to the users regarding the recommended item [13]. An example of this can be based on content-based recommendation "*since you liked the book x, we recommend book y from the same authors*" [13].

The explanations for constraint-based recommendations have to explain to the user about the constraints applied by them. These explanations can contain additional information apart from basic information. For example, *"since you prefer taking sports photos, we recommend camera y because it supports 10 pics/sec in full-frame resolution. z would have been the other option but we propose y since you preferred purchasing from provider k in the past and y is only a little bit more expensive than its competitors"* [12]. This explanation will help the user understand the constraints they have applied and how they can choose other competitor's items as well. This could be an example of the goal "transparency".

Recommendations that are based on the knowledge of the domain, have explanations that will explain the criticism it wants to convey about the products recommended. These explanations are called critiquing-based explanations. This could be done for deep item knowledge understanding. For example, *"item y would be a good choice since it is similar to the already presented item x and has the requested higher frame rate (pics/sec)"* [12]. It is also possible to generate comparative explanations [8] in the field of tourism. The authors have reduced the overhead of remembering the names of the places and descriptions by displaying the explanations side by side. This is one way of critiquing based explanations. This could be an example of the goal "transparency" with *additional* information.

Finally, the "no solution found" issues can appear due to the extreme constraints that can be applied by the user. The explanations will give ideas on how to go around the situation by suggesting solutions. For example, *"if you increase the maximum acceptable price or decrease the minimum acceptable resolution, a corresponding solution can be identified"* [12]. This could be an example of the goal "Satisfaction". Additionally, we can mention about additional information that is necessary to solve the issue of "no solution found". This additional information can contain "asymmetrical dominance" (comparing not just price but also quality for example) or "compound critics" (compares items explaining about resolution, price, and quality in the domain of camera).

In the next section, we will discuss how explanations are adapted to groups and how explanations can be explained for sequences of recommended items.

2.2.3 Explanation for Groups

The related work on designing explanations for groups are discussed here. Firstly we will discuss the goals that group decision making has to consider when designing explanations for the groups. Secondly, we see some related work that makes use of these goals that are suitable for particular aggregation functions for groups. Thirdly we see how individual explanations can be extended to traditional group recommendation algorithms. The textual explanations explained in [12] have complete information or have preserved privacy.

2.2.3.1 Goals for Group Explanation

An explanation designed for groups has additional goals to the ones mentioned in Section 2.2.1. These goals are *Fairness, Consensus and Optimality* [12, 40]. Firstly, Fairness is to take into consideration, the preferences of all group members as far as possible, secondly, Optimality is to find the most favorable or near optimum solution, finally, Consensus makes the group members agree on a common decision.

For example, travelers who have to agree on a general agreement would like to know *"why"* an item has been recommended to the whole group when every member would have a difference in opinion. The individual group members in this manner would assess the extent to which the recommender system has taken their preferences to make the recommendation. Also, in contrast to explanations (section 2.2.2) given to individuals explicitly, group explanations have to stress on the aspect of how "the interests of individual group members are taken into account" [12].

The aggregated preference approach supports the goal of fairness and consensus whereas the aggregated models show explanations in terms of group-level [12, 35]. Explanations that have been discussed in section 2.2.2, can be adapted to groups as well. "The explanation given for groups are affected by the underlying recommendation algorithm" [12]. As we discussed in Section 2.1.1, aggregated preferences and aggregated model can be applied to the traditional recommendation algorithm.

Tran et al [47] have designed and generated explanations (of aggregated preference type - see Section 2.1.1), that consider "preferences of all" or the "preferences of the majority of group members". These types of explanations give good results when tested for the group aspects of "*consensus*" and "*fairness*". They also found there was a positive correlation between explanations with perceived consensus (fairness) and satisfaction for the given group recommendation. They have achieved this conclusion by testing three types of explanations. Here basic explanation describing aggregated preferences is type 1 and type 2 and 3 have decision history and future decision plan additionally. For example, *"Item X has been recommended to the group since it achieves the highest total rating"* [47] is the basic explanation. This explanation is concatenated with either decision history or future decision plans that consider the fairness of the priority to make the next decision.

Najafian et al [12] has designed a "repair-related" explanation to describe user's aggregated preferences with pleasure as a basis and fairness as a basis to reach acceptable "*consensus*". The authors designed "repair-related" explanations which repair the "inconsistency" in the individual user opinion when compared with the group. The explanation they designed is *"Item y is recommended because nobody hates it in the group due to the lowest rating determined for the user a and supports the highest rating determined for user b"*. As we notice, the explanation is dependent on the underlying aggregation preference function and is designed with "consensus" as the group explanation goal. They have also designed explanations with "fairness" as a basis. It would be as follows *"The system detected you might not like song 1 but it is the song Mary prefers most. You made your choice in the previous round, now it's Mary's turn"*. Another example, of how aggregated predictions are explained *"one of the group members has specified the lowest maximum price of 500"* [12].

2.2.3.2 Group Explanations dependent on the aggregation strategies and recommendation algorithm

The group explanation styles can be based on the (1) aggregation approach i.e. aggregated predictions or aggregated model, (2) hybrid (mix of both aggregation types), (3) no solution found and (4) group's social reality [12]. The explanations designed for groups could be affected by the applied social choice strategies discussed in Section 2.1.1.3. Also, the individual explanations given for traditional recommendation algorithms can be extended as explanations for group recommendations [12]. The authors of [12] have categorized visual and textual explanations in terms of the underlying recommendation algorithm. They have presented explanations for aggregated models and aggregated predictions for each type of recommendation algorithm.

In ***collaborative filtering*** explanations such as *"users who purchased item x also purchased item y"* [13] or *"groups who purchased item x also purchased item y"* can be used to explain group recommendation done after applying collaborative filtering. When the social aggregation strategies are applied for aggregating the ratings, the explanation for the recommended item y given for group can be explained either as "*item y has a group score of 2.9 due to the (lowest) rating determined for user a*" or as "*item y is recommended because it avoids misery within the group*" [12]. The former is based on aggregated predictions where the user responsible for the lowest score is reported and the later for an aggregated model which preserves privacy designed especially for the "Least Misery Strategy" aggregation strategy. This kind of explanation can be extended to other strategies discussed in section 2.1.1.3. The authors of [12] have extended the work of [20]' s histogram and made "frequency distribution" for displaying the nearest group's ratings or nearest neighbor's rating. Also, they have suggested "spider diagrams" for understanding each group member's ratings on every item to make a group decision.

In ***content-based filtering*** the content similarity between item descriptions and the relevant keywords present in the user profile are used to predict items. Similarly, social aggregation strategies can be applied to this recommendation to generate group recommendations. For example, the explanation given to the group while it is aggregated prediction is "*item t1 is recommended since each group member is interested in category cat2*" [12] or when it is the case of aggregated models, privacy is preserved by the explanation, "*item t1 is recommended since the group as a whole is interested in category cat2*" [12]. These explanations can be represented most effectively by "TagClouds" for explaining the recommendations made for the group [48]. Here most words are shown by changing font and colors depending on their frequency for every user.

In ***constraint-based filtering*** the domain knowledge of items and its constraints help recommend items. The explanations designed answer "how", "why and "why not" (no solution found) using the constraints given by the users [12]. The relevant user requirements that were chosen to generate group recommendations are used to explain recommendations. The authors of [12] have designed the relevant explanations for different situations arising due to constraints of users. The situations were the (1) need for combining the use and constraints of items for groups, (2) need for explaining the situations where the individual users are inconsistent with each other's preference and to invoke a consensus by adapting their own preferences [35], (3) need for user-generated explanations by the decision task creator, to explain the constraints chosen to make the final decision, (4) need for fairness in deciding the items during repeated group decisions.

In ***Critiquing-based*** group recommendation and explanations, an explanations from the work of [12] such as, "*the price of camera t1 (299) is clearly within the limits specified by the group members. As expected, it has an exchangeable lens. It has a resolution (24) that satisfies the requirements of u1 and u2, however, u3 has to accept minor drawbacks. Furthermore, the weight of the camera (1.5) is significantly higher than expected by u1 and u3*", clearly explains the criticism and the requirements of every user in the group.

These above-discussed explanations in this section, address neither the social reality of the group like personality, tie strength, etc., in detail nor have explained about the order of sequences that were generated due to the aggregation strategies. This will be seen in detail in Section 2.3.2. Before we move on to the social explanations and group attributes, let us first see the other relevant group explanations that do not depend on the underlying recommendation algorithm.

2.2.3.3 Other ways of Explaining for Groups

In the previous section, we have discussed explanations for groups that depend on the aggregation strategy. In this section, let us see in brief about other ways of making group explanations, that does not necessarily depend on the aggregation algorithm.

Demographic filters were applied in the making of INTRIGUE a mobile and web-based tourist itinerary recommender. The system [3] generates a tour schedule for heterogeneous groups. Explanation generated after recommending the itinerary is as follows: *"For children, it is much eye-catching, it requires low background knowledge... For yourself, it is much eye-catching and it has high historical value. For the impaired ..."* . The words like children, impaired contribute to demographic information.

The tourist trip design problem can be solved using RS [21]. They have used Dijkstra's algorithm, its extension, and the Greedy randomized adaptive search procedure to generate a sequence of the place of interest (POI) recommendations to the group. But they have not used explanations to support the recommendations.

The authors of [36] have made recommendations and explanations to support group decision making. The explanations display the number of persons who liked a POI (for example if there are 3 members and only 2 members like a place, then "2/3 likes this" is displayed). The explanations also display the user's current state of physical energy, time of the day, rating preference of members and content similarity to convince other users about the places recommended.

An example of interactive recommendation and explanation for recommending POI for groups is "Travel Discussion Forum" developed by the authors of [25]. The vacation preferences of the group members are specified collaboratively. The similar group members are chosen based on similar item preference and these users arrive on a *consensus*. The interface gives the "option to copy or view" other group member's preferences; animated characters representing group members who are not available for discussion or communication use speech, facial expression, and gesture to arrive at a consensus. There are two aspects the visual interface creates awareness of other people's preferences and responses. It can be either in graphical representation or human-like representation.

2.2.4 Summary

We have seen explanations that are designed for groups in the domain of art, travel, online shopping, movies and books in the above sections. Firstly we have discussed the goals of the group explanations, secondly, examples of explanations for groups that depend on the aggregation strategies. Thirdly on explanations for groups that do not use aggregations strategies. The authors of [12] have quoted that, "Different ways to explain group recommendations depending on the used aggregation function(s) are an issue for future research". This suggests that we should try different ways of formulation of explanation for the underlying aggregation function. This has been discussed in Chapter 3 particularly in the Section 3.2.

The explanations discussed here in this section for groups do not involve group attributes. This research regarding explanations involving group attributes are is not covered in [12]. Yet, the authors have discussed group attributes and the incorporation of the same group recommendations. Explanations that involve group attributes like relationship strength, personality, emotions, conformity, satisfaction factors make us understand that explanations and recommendations cannot be solely based on group member's rating [38, 37, 16, 31, 13]. We will discuss the importance of considering group dynamics in recommendations and explanations in the next Section 2.3.

2.3 Group Attributes and Explanations

The group member's rating preferences for item(s) or individual satisfaction with the recommended item(s) can be affected by the group's dynamics like relationship strength between individuals or a member's personality. These social interactions are external to RS. Current group recommendations assume that there is no rating (item) preference dependence and hence do not consider group attributes like social relationships or interaction types among group members while recommending item(s) [46, 13].

Consequently, these group dynamics (group's attributes) also affect an individual's satisfaction with the recommended item(s).

Individual's Satisfaction with the recommended item(s): The individual's satisfaction with the recommended item in a sequence can be dependent on the previously recommended item [32] or dependent on other group's overall satisfaction [31]. According to social psychology, individual satisfaction is affected by two aspects of the group [31] They are (1) Effect of other group member's emotions on an individual's satisfaction (*Emotional Contagion*), (2) Effect of other group member's opinions on an individual's satisfaction (*Conformity*) [31, 13].

In emotional contagion, the individual's satisfaction can depend on the personality of other group members or relationship the user has with other group members. The relationship type an individual has with group members can affect the emotions of an individual group member. The relationship type that affects emotions is by the order of close friends and people whom they respect than with people with whom they feel are equal and people with whom they compete [31]. In conformity, a strong personality's satisfaction cannot be easily affected by other group members' opinions.

As we see a person's personality or relationship strength seems to play a role in the individual's satisfaction with the recommended item. In [32] these social aspects are incorporated in their satisfaction model. Alternatively, these aspects can be used while choosing an aggregation strategy [31].

2.3.1 Group Attributes

We discussed the aspects that affect an individual's satisfaction. In this section firstly we will discuss group attributes (group dynamics) and the way they can be applied to group RS [13, 31]. Secondly, we will discuss some examples that use these attributes while recommending an item or a sequence of items to a group. Finally, we will discuss how relationship strength could affect the individual satisfaction for the given explanation that based is on underlying aggregation strategy as a research gap.

The authors of [31] have classified the group's attributes that influence a group member's decision of accepting a final recommendation or giving their final opinion on a candidate item (for example rating). Hence, group attributes are important to understand and implement in group recommendation and eventually use them in explanations to improve the explanation's goal.

The group attributes are classified as (1) Roles that people play, (2) Personality of users, (3) Expertise, (4) Relationship Strength, (5) Relationship Type, (6) Personal Impact [31]. The authors of [31] have further grouped the above-mentioned attributes (1 to 6) based on how they have been incorporated by group recommenders for recommending items. The further categorization listed is: (*Type 1*) based on the **individual** group attributes used ,particular group members can be assigned more weight-age (for example expertise in selecting a movie to watch, books to read), (*Type 2*) group attribute as a **whole** is used on the group to decide its impact on satisfaction and group decision (for example relationship strength among group members) and finally (*Type 3*) the attributes can be used to adjust the ratings based on the group attributes that individual share with the other group member in terms of **pairs** (for example making decisions with group member you respect).

Type 1 In the work of [3], the group members are grouped based on their demographic roles like children, adults with and without disabilities. They have assigned more weight-age to children and people with disabilities while recommending a place of interest to the group. The authors of [3] have additionally given explanations based on these demographic styles. Another work is from [41] where "dictatorship" is used to assign weights to individuals and then recommend items to the group. These are some examples of group attribute "*Roles that people play*". The work [37] for their application "Happy Movie" have assigned weights to individuals based on their *personality*. In [16], they have assigned weights to people who have *expertise* in the field of movies while recommending movies.

Type 2 The work of [16], where the group attribute of relationship strength has been considered as a whole. The results of this work were that (a) person who shares a close relationship (couples or close friends) should be subjected to the strategy of "Most Pleasure". The groups that share a strong relationship is called **primary groups** [34] (b) people who share no personal relationship otherwise called **secondary groups** [34] should be subjected to Least Misery Strategy, (c) people who are in the middle of these two extreme relationship strength (intermediate) should be subjected to average strategy.

Type 3 In the work of [37], where the group attribute is included in the aggregation strategy to adjust the rating of the individual based on the relationship strength with other group members.

Therefore in this section, we have seen the different types of group attributes and how they are influencing the group recommendation. Yet this section has not discussed how the group members have accepted the group's recommendations when these group attributes are incorporated into the explanations. This is discussed in Section 2.3.2.

2.3.2 Explanations involving Group Attributes

In this section, we will discuss some examples of how group attributes are used in the explanations and how the group members have received this kind of social explanation.

The existing work of [37] has used the group attributes of personality, tie-strength, and satisfaction value of each member with the given item to improve their group recommendation. The results showed that the presence of information about other group members' tie-strength and personality together, made good progress on the group recommendation's acceptation in contrast to using just one of the said group attributes.

Quijano et al.,(2017) [38] have extended [37]'s work and generated explanations that separately were designed for each group attribute (personality, tie-strength, and satisfaction) and then presented one by one based on the user's individual rating for each recommended item. They call these explanations "personalized social explanations". This is because the explanations for personality conveyed how the other person would accept other people's preferences by leveraging social interactions.

For example, "*Although we have detected that your preference for this item is not very high, your friends X and Y really like it. Besides, we have detected that they usually don't give in*" [38]. This explanation is given for explaining rating preferences. For the explanation that conveys the strength of relationship to increase the acceptance of the group recommendation, the authors gave an explanation that invoked "social bonds. The explanation was *"Although we have detected that your preference for this item is not very high, your close friend X (who you highly trust) thinks it is a very good choice"* [38]. Finally, to display how the satisfaction value is for the other group members and increase the acceptation of the group recommendation, the explanation displayed was *"Last time users X and Y gave in with the selected choice, it would be fair if this time they were given some kind of priority"* [38]. This work has suggested how to adapt explanations to group member's relationship strength, personality, and satisfaction but there is no suggestion on how to adapt explanations without social component to the group's relationship strength.

Quijano et al., [38] have designed social explanations that invoke social bonds relating to relationship strength to increase individual satisfaction. They also bring to effect the "consensus" part for the group by informing them about rating preference and personality of other group members. This brings us to the research question of whether groups of different relationship strength (primary and secondary) accept explanations (that do not invoke social bonds), that explain about other group member's preferences with their respective names and ratings by being system transparent.

Delic et al., [9] have improved the group RS made for travel by incorporating social relationships strength as the group attribute. The authors have results that indicate that a "socially central" person was not influencing the decisions inside a group of people who were considered equal to each other. They also found the groups that are closely related to each other (primary groups [34]) had more satisfaction with the group's decision than groups who did not share a personal relationship with each other (secondary groups [34]). The authors have also concluded that *social relationships* can be used to predict the overall satisfaction of the group on the choices made by the group and every member's individual satisfaction with the choice made by the group. This work has suggested future work on choosing a social preference aggregation strategy for the group type (close or non-close).

Gartell et al., [16] have improved group recommendation by incorporating the attribute of expertise and social relationship strength into the group recommender system. They conducted experiments with participants using the movie domain. The frequency of contact between users helped them decide the strength of the social relationship interaction. The results showed that different aggregation strategies - most pleasure, least misery, average - were best fit for different group strengths - strong, weak, intermediate - respectively. These authors have not designed explanations that would satisfy the users with particular relationship strength but in contrast, improved the group RS by switching between strategies to satisfy groups that fit different relationship strength.

Delic et al., [10], have observed groups of people who were on a discussion for deciding on a trip to ten predefined places. They have made some conclusions. They were: the individual's satisfaction depended on the process in which a decision was made, the characteristics of the group member and the group as a whole. They did an analysis by comparing the group choices generated by aggregation strategy with the individual's "pre-discussion" ratings. They performed this analysis for multiplicative, least-misery, most-pleasure, median and additive. The best performing strategy was multiplicative. Even though the multiplicative strategy was one of the best performing strategies, we decided to go ahead with the work of [35] and choose an aggregating strategy that combines least-misery, average without misery and most pleasure to generate group choices of places of interest as a sequence. This is because this strategy works well with tourist RS and produces a sequence [27].

Kapcak et al., [27], gave an initial template to the crowd and then improved them. They had an explanation template based on predefined situations. The templates looked like (a) "*Hello John, we know you would love to see the Eiffel Tower, however, others in your group would love to see the Louvre first*" [46], (b) "*Even though you wanted to visit POI X, most of your friends gave a very low rating for that POI. Therefore, we did not include that into the recommended POI's for the group*" [46]. They had generated explanations based on the aggregated predictions for an individual item in the sequence using synthetic ratings (from [29]) with the aim of consensus [35]. The authors have suggested that the Thomas-Kilmann Conflict Mode Instrument's (TKI) conflict style as a personality model or relationship type can be considered in aggregation algorithms and as well as explanations as future improvements.

In the next section, we have discussed how the designed explanations are evaluated by various metrics based on the goal they would like to achieve for the group RS.

2.3.3 Evaluating Explanations

The effect that explanations have on users has been evaluated by experiments involving user studies throughout literature. In this section, we have discussed various evaluation approaches, metrics and the effect that explanations with certain goals for groups have achieved. The examples have been discussed below.

Quijano et al., [38], have conducted user studies to validate if textual or graphical explanation designed by them was preferred over the other by the users of the "Happy-Movie" application. These explanations did not preserve the privacy of the individual members in terms of the names and preferences of group members. They had additionally included the social components of *personality, tie strength and previous user satisfaction* with the recommended item in the explanation. These explanations were given in a textual format and graphical format to the users. The goal or the metric that the explanations were trying to improve or achieve was to increase the "Happy-Movie" application's persuasiveness, usage, efficiency, trustworthiness, and individual and group satisfaction with the system's recommendation. These goals were evaluated by presenting a questionnaire to the participants of the study who evaluated the explanations by answering the questions (with five stars Likert scale) in the *questionnaire*. The results were analyzed for statistical significance using the "Kruskal-Wallis H test" and "pairwise Wilcoxon test" to prove their hypothesis (of including all social components in the explanations). Kapcak et al., [27], used a "crowd-sourcing technique to generate and validate the explanations from an initial template of explanation. These explanations were designed for individual users explaining the preferences of group members in the domain of tourism. The metrics that the crowd-workers used to evaluate the explanations are "Quality" (informative), "Quantity"(truthful) and "Relevance" (relevant information to the social scenario presented). The explanation was not designed for the sequence not the privacy of the users were preserved.

Najafian et.al., [35], have created two types of explanations namely repair-related and reassuring explanations with both complete and vital information. These explanations are transparent as they have explained how the ratings are aggregated. The authors have evaluated the explanation's ability to increase user satisfaction by structured interviews. the questionnaire of the structured interviews had a 5-point Likert scale question on the satisfaction level, additional questions of what the user liked about the explanation and what the user would like to change about the explanation. These questions helped them calculated the average satisfaction of the users to decide which explanation type best fit the five scenarios designed by the for a sequence of recommended songs. These sequence were decided based on the sequence of songs recommended. This sequence's order can cause some disagreement amongst the group members. For example, the scenarios in which the song not like by a member is being played, the song liked by a member has not been selected yet to be played nor has never been recommended at all. There can be situations in which the group agrees on the recommended song unanimously.

2.3.4 Summary

Firstly, we discussed the explanations based purely on the aggregation strategy (see Section 2.2.3.2, 2.1.1) for groups. Secondly, we discussed the explanations that were dependent on other underlying recommendation algorithms in Section 2.2.3.3. Thirdly, group attributes are explained in Section 2.3.1. Fourthly, we acknowledged how group attributes are incorporated into the group RS to improve group recommendations and also generate explanations based on these group attributes (see Section 2.3.2). Finally, we discussed the evaluation of explanations.

2.4 Discussion

There has been much work in recent literature in incorporating group attributes to improve group RS. These approaches involve one or a combination of social interactions such as relationship strength, personality, or the expertise of group members, to improve the way aggregation strategies were recommending sequences or single items to a group. For example, these approaches incorporated the relationship strength present amongst group members directly [9] or indirectly [16, 37] into the group recommendation algorithm.

A similar shift in research focus from direct explanations explaining the group's preferences to a group member [35, 27, 15] to explanations that involve group attributes [3, 37] (also known as social explanations) has been found in literature. Social explanations usually invoke social bonds or the strength of the personality of group members while explaining to individual users about the group's preferences. This aids in convincing members to create consensus [38].

We have seen how a high-level of transparency helps improve user satisfaction with the recommended item(s) as transparency helps understand the other group member's preferences [27, 35, 12]. The work of [27] has also acknowledged the importance of privacy when the relationship strength among group members is considered while explaining a recommendation. For example, with acquaintances, the group members may not feel comfortable in revealing their names and specific preferences about an item. There is acknowledgment in the literature to focus on evaluating whether transparency or privacy is preferred in an explanation (without social component) for different relationship strengths [27].

There has been limited work in comparing which explanation style is best to improve user satisfaction regarding the group recommendations [35, 47]. In [18]'s work, the authors have compared different explanation styles and evaluated their effect on user satisfaction. They have concluded that perceived transparency highly contributed to long term user satisfaction and trust in the RS compared to effective explanations. In [47]'s work the authors have proposed different explanation types for different social choice aggregation strategies. They have investigated which explanation type performs best to satisfy the goals of a group explanation (consensus, fairness, satisfaction). They found that explanations that consider all or most of the member's preferences performed the best in improving satisfaction. Yet, there has been no work that compares which explanation type is best suited for different relationship strengths to increase the user's satisfaction. In [23]'s work, they experimented by displaying recommended tourist trips via three means (public display, mobile phones and a distributed system which had a combination of both). The results showed that group members who were open to a discussion were comfortable with a public display but that was not the same case for group members who did not know each other. The latter type (considered as acquaintances) were comfortable in displaying only "selected content" on the public display after they rated places they wanted to visit. From this work we can infer that acquaintances would prefer some privacy compared to a group where members knew each other well.

2.5 Research Gap

The gaps identified after the literature survey are listed below.

- An explanation for the entire sequence of items such as an itinerary (in the domain of tourism) has not been well-reported. This is important for tourists who need a holistic idea about their recommended itinerary.

  - In the domain of music, satisfaction-oriented explanations with the aim of creating consensus were designed for individual items in the sequence. The explanations contained information about the inconsistency of preferences of other group members [35]. These explanations were given for each scenario where disagreement with the order of the sequence can occur. However, the whole sequence was not explained.

  - In the domain of tourism, satisfaction-oriented explanations that invoked social bonds (such as "friends") with the aim of consensus were designed for each item [27]. For example, it also expressed other members' rating preferences to explain why a POI was not in the preferred order in the sequence given. Similarly, the whole sequence was not explained.

- Conflicting situations can arise due to disagreements with the order of the recommended sequence. These situations can be explained to the user. There is limited or no work in designing an explanation for the complete sequence with the **possible conflicting situations**.

  - Explanations have been designed for *individual items* present in a sequence [35, 27, 12]. These explanations are designed for just one conflict situation at a time and do not explain all the possible situations that could arise.

- Explanations are designed by displaying the item name, group member's name and the respective rating given by that person for the particular item. However, designing explanations that consider user's **privacy** while explaining group preferences to an individual is not studied extensively.

  - In the work of [12], aggregated group models reduced privacy concerns of the group members by recommending to a group profile and explaining them. For example, "A majority think that it is a good choice. Some group members think that it is an excellent choice" [12]. Yet a similar method has not been approached in explaining aggregated predictions.

- There is limited or no work in **comparing** Transparent and Privacy-Preserving explanations to check which type increases user's satisfaction based on the relationship strengths within a group.

  - There is work that acknowledges the privacy demanded depending on the group type (acquaintances or couples/friends) [27, 23]. There was no explicit comparison made to determine the preference of privacy for groups with different relationship strengths.

2.5.1 Thesis Motivation

In this thesis, we aim to design an explanation for the entire sequence at first. Then we *evaluate* whether the designed explanation (or privacy preserved variant of the explanation) influences user satisfaction based on the different relationship strengths within a group. We achieve this by doing the following.

1. Formulate explanation for a sequence of POI focusing on user satisfaction (Preliminary Study 1).

   - Possible conflicting situations that may arise due to the disagreement with the sequence order for each group member are considered.
   - Then we perform qualitative analysis on the user comments and investigate the uniform textual content to be presented in the explanation for each social situation.

   Motivation : The motivation for investigating the content of an explanation is inspired by different versions of explanations that we saw from the literature. For example, Kapack et al., [27], used user friendly words in their explanation such as *".. we know you would love to see Tour Eiffel, however others in your group would love to see Louvre first"*, to convey preference of other group members to a group member. Najafian et al., [35] explained the preferences either by mentioning the person's name who was responsible for the lowest and highest rating and their corresponding rating value. Felfernig et al.,[15] has explained the preferences to group members using names and real number values such as *"since the maximum camera price accepted by group members is 500 (defined by Paul) and the minimum accepted resolution is 18 mpix (defined by Joe), we recommend y which supports 20 mpix at a price of 459"*; Quijano et al., [37] has used social explanations to explain the preference of other group members such as *"Hi there Jaime, we have predicted that you will be just basically okay with this movie. However, we have detected that your friend Claire, who you are really close to and trust, will love it."* Therefore, we see that there are many ways to present the information about item rating, item name and the person responsible for it. We wanted to investigate some of the best ways of presenting explanations that was satisfactory to users.

2. Formulate and structure explanations for different relationship strengths

    - The need to change the formulated content according to the relationship strength is studied.
    - The structure of the explanation for the entire recommended sequence is studied. We also study the order in which information could be presented according to the effect created by the social situations.

    Motivation: A constructive way of giving feedback to people is discussed in [11]. The authors analyzed that giving only negative feedback can induce strong emotions and the person receiving it can reject it. They suggest giving negative criticism in-between two positive criticism. This type of giving feedback is called a "feedback sandwich" [11]. An empirical insight is that the user can be dissatisfied when their favorite places are not in the preferred order or not recommended at all or when a place they dislike is being recommended. The explanation for the entire recommended sequence is a combination of such facts about the disagreement when it is about the order of the sequence. Hence conveying such facts at a suitable position in the textual explanation will help the user to accept the recommendations and understand other member's preferences. This may also keep the relationship between RS and user healthy.

    When people do not share a personal bond (secondary group), they may appreciate some privacy when rating preferences of group members are explained by the explanations given by the RS [27, 23]. Hence we would like to explore this need for privacy.

3. Evaluate whether different explanations formulated for different group types influence an individual's satisfaction

    - The two group types are primary group (couples/close-friends) and secondary group (acquaintances).
    - The two explanation types are Transparent and Privacy-Preserving for aggregated predictions for the group members.

    Motivation: In literature, [27, 13] there has been attention to the research gap where group members may feel a breach of privacy when they traveled with people they did not know. The explanations may actually create strife than create consensus by revealing personally identifying information such as names. We would like to investigate if this is really the case by conducting a user study and evaluating how the satisfaction levels were affected.

Chapter 3

Preliminary Studies

3.1 Introduction

After identifying the research gap in the literature study, the three research questions are: (1) How should one formulate explanations for user satisfaction? (2) How should one formulate and structure an explanation for different relationship strength? (3) Do different explanation types for different relationship strength influence user satisfaction?

To answer the research questions, we have conducted three user studies in total. *User studies 1 and 2* are preliminary studies found in this chapter and they are *exploratory*. User study 3 is found in Chapter 4. The first study will investigate suitable textual content to be presented in the explanation. The second study explores the structure of the explanation for the entire recommended sequence and if there a need to change the formulated content according to the relationship strength. The final study investigates how the transparent and privacy-preserving explanation given for the entire sequence (possible social situations combined) affected an individual user's satisfaction when they belonged to groups with different relationship strengths (close friends or acquaintances).

Motivation for preliminary studies: The recommended sequence is an itinerary which has to be visited in the recommended order by all the group members together. As we discussed in the literature (Chapter 2) there are conflicting situations that may arise due to the disagreement with the recommended (POI) sequence's order. These situations may or may not occur in a group, based on the sequence order that the chosen recommendation algorithm has generated. The disagreement situations that every group member could face due to the sequence order are:

- Places a user likes (rated high) is not recommended initially in the sequence
- Places a user likes (rated high) is never recommended at all
- Places a user dislikes (rated moderately or low) is recommended.

Also, the positive situation where the group member generally agrees on the order and do not have a disagreement is,

- Places a user likes (rated high) is recommended

For the **first preliminary study**, the literature motivated us to investigate the various ways of presenting explanations to an individual group member. Therefore, for each of the four possible social situations, we have proposed an average of three explanations. These explanations focus on improving satisfaction of each group member with the given explanation's content. Finally, the user's comments given for all explanations are collectively analyzed. Then a uniform textual content to be presented based on each social situation is decided.

For the **second preliminary study**, just one social situation is chosen. Then the explanation determined at the end of the previous study is proposed to the users (n = 12). The social situation chosen is "Places a user likes (rated high) is not recommended in the preferred order". Then the explanation is given to both the group types. After the user study, the user comments are analyzed with three different aims in mind. They are: improving the textual content determined in the first preliminary study, retrieve useful comments about the sentence order and finally, if there is a need for privacy requested by group members of both the group types. The collective comments from study 2 will help us design an explanation for the entire recommended sequence of POIs accommodating the possible social situations with uniform content. Also, it will help us understand if there is a need for privacy by the groups.

The **final user study** (user study 3) will evaluate if the two types of explanations (transparent and privacy-preserving) influence the individual satisfaction of the group members present in groups of different relationship strengths.

	A	B	C	D	E	F	G	H	I	J
John	10	4	3	6	10	9	6	8	10	8
Adam	1	9	8	9	7	9	6	9	3	8
Mary	10	5	2	7	9	8	5	6	7	6
LM	-	4	-	6	7	8	5	6	-	6
MP	-	9	-	9	10	9	6	9	-	8
Sum	-	13	-	15	17	17	11	15	-	14

Group List: (E, F), (H, D), J, B, G (threshold 4)

Figure 3.1: Synthetic ratings and Recommended sequence using a hybrid aggregation strategy [29, 35]

Social strategies chosen to generate a sequence: All the three studies were controlled by using synthetic ratings (Figure 3.1) for 10 artificial places A until J in alphabetical order. These ratings can be the result of any traditional recommendation algorithm discussed in Chapter 2 (For example: collaborative filtering or content based filtering etc.,). The places are not named by real place of interest names (such as "Van Gogh museum" or "Tour Eiffel") as it could have an effect on the user's satisfaction during the study which is not in the scope of this thesis. These places are rated by a group of 3 users (John, Mary and Adam). This thesis has not considered the effect of explanations for groups more than 3 users for the sake of clarity in this study.

The aggregation algorithm chosen has applied three strategies on the ratings one after the other. The strategies are most pleasure, least misery and average without misery [35, 29]. Hence the sequence of places generated would be: "(E, F), (H, D), J, B, G". With this information, we have chosen the four possible social situations and, designed explanations for it. The rest of the sections are about preliminary study 1 and 2's study design, procedure, results, and discussion respectively.

3.2 Preliminary Study 1 : Content

The research question focused through this study is: *How should one formulate explanations for user satisfaction?* Therefore, various variants of satisfaction-oriented explanations are investigated.

3.2.1 Materials for the study

The explanations examined the content that focused on expressing other group members' preferences and ratings to an individual group member about the recommended items in the sequence.

Content of the variants of Explanations shown: The explanations can be thought of as a combination of various information blocks that represents the reasons for the recommended item. The blocks were: (1) Information on algorithmic transparency, (2) Information on system's transparency describing ratings of users, (3) Information describing the situation the user is facing. The various variants of the explanations conveyed the same meaning about group's preferences for a social situation under discussion. The content in these blocks *differed* in terms of:

- Various words choices or phrases were used to convey the information on algorithmic transparency, rating preferences of group members and social bonds present among group members. For example,
  - Either displaying the rating value of group members as *real numbers* (1 being the lowest value given to a place and 10 being the highest on a 10-point Likert scale) or as *categorical values* ("low" depicting rating values [1,2,3,4] , "moderately" depicting [5,6] and "highly" depicting [10,9,8,7]). The segmentation of rating values into categorical values are chosen empirically.
  - Mentioning the ratings explicitly with a combination of names and rating of the group members involved. For example, *"Mary rated place B as 5.."* or *"Mary rated place B moderately"*
  - Usage of the social term "friend", "best friend", "close friend", "friend whom you trust the most" in the explanation to invoke social bonds. This was done while explaining about the rating preference of each member.
- Position of the above said word/phrases in the explanation
- For explaining how the system/algorithm works, we chose,

- To vary the phrases that use user friendly words or technical words
- Explain the algorithm chosen as the reason for a favourite item to receive low priority in the sequence order; or when justifying disagreement situations

Therefore, twelve variants of the explanations, across all the aforementioned situations were created. These were a combination of the factors listed above.

Sample variants of Explanations shown: Based on the chaining of aggregation strategies (most pleasure MP, least misery LM and Average without misery (threshold)), as shown in the figure 3.1 a sequence is generated. Few sample variants of the explanations for social situations is shown below. The content present in the explanations with variants could have an effect on user satisfaction and some of them are shown below to understand how the explanations varied from each other.

- Dear Adam, you are not recommended some of your favorite places B and J which you have rated 9 and 8, first in the trip. This is because your close friends Mary and John whom you highly trust have rated only 5 and 4 for B and 6 and 8 for J. They would be unhappy if these places are chosen first. However, you will visit these places you like later during your trip.

- Dear Adam, even though some of your friends Mary and John's top choices E, F do not match with your top choices, they are recommended first and you can visit your favorite place of interest B and F later in the trip.

- Dear John, the place of interest A and I are not favored by Mary and Adam. This is because these places are rated low by one of the group member and hence they are not included in the final recommended sequence. Other group members have agreed to *compromise* on their favorite place C to keep you happy.

- Dear John, the places of interest A and I are not favored by your best friends. Adam has rated A as 1 and Mary rated I as 3. This system supports ratings that causes more pleasure and avoids places that are heavily disliked by other group members. Moreover, your close friend Adam does not get to visit C as he is accepting your choice of disapproval of the place.

- Dear John, the place of interest B is being visited now even if you have rated it low. Adam and Mary have similar preferences as yours on the item recommended. They have agreed to visit this place to keep other group members happy. You would have visited some of the places you love prior to this recommendation.

Section 3.2.4 discusses the qualitative analysis that was made by observing user opinions of preliminary study one.

3.2.2 Study Design

Different variants of the explanations that conveyed the similar meaning were created for preliminary study one. This was repeated for each of the social situations. The study had five participants. All were students of TU Delft varying between the ages of 23 to 26. They were 40% male and 60 % female. They were **not briefed** neither about group types nor about the explanation types. They were shown a set of explanations for every social situation and asked about their opinion of what they liked and disliked about the explanations shown. The reason for this low number of participants is because we are not constructing entirely novel explanations, but adapting explanations from literature. The descriptions on the content used in the explanations have been discussed in Section 3.2.1.

3.2.3 Procedure

To acquire different user opinions and make a decision on the content that could give more relevant information about group member's preferences, this study is performed via a questionnaire. The open *questions asked* to the users for every variant of the explanation presented were:

- What do you like about this explanation?
- What would you like to change about the explanation?

The participants' opinions on the questions would help us understand the reasons behind likes and dislikes expressed by them for the content displayed in every variant.

3.2.4 Results

The following **analyses** were made from the likes and dislikes expressed by the participants as user comments about the explanations. These analyses contributed to the suitable content of the explanation.

1. *"The influence of strong words/phrases derails the focus of the explanation (which is to provide satisfaction by providing algorithmic transparency and information about group member's rating preferences)."*

    - The words like "..the top choices.." gave a creative look but was not conveying the synonym of "favorite places".
    - The phrases of"..your choice of disapproval..", "..rated extremely low..", ".. unhappy.." to name a few have induced a negative affect on users than conveying the synonym of "the rating of a place".
    - A few phrases like "..to keep you happy..", "..keep the satisfaction high by excluding items.." made users feel they were creating havoc in the group (primary/secondary) due to the presence of their personal preferences.

    Therefore the usage of such strong words when designing the explanation could be reduced.

2. *"Need for categorizing rating values into low-medium-high scale."*
   The absolute numbers present in the explanation created an unnecessary overhead of remembering numbers for the users.

   - For example "..place B and D were rated 5 and 8 respectively.." made the explanation too machine-like.
   - This posed a further problem of a messy look and *explosion* of the length of the sentence when there were more than three users or there were lots of other ratings to explain for instance.

   Therefore an anonymous scale of "low", "moderately" and "high" for the rating values can keep the user informed of the group member's rating preferences at the same time make it easy for them to read and understand the explanation.

3. *"To let the users know that they are not the only person in the group to receive recommendations with a low rating."*
   This preliminary study also gave the intuition to improve on some pointers that were not present in the variants given.

   - Introduce the information that will explain to the user that they were not the only person "compromising" for the entire trip in the next study. This could help in increasing the understanding level of the users regarding the group's preferences.

   This will in-turn contribute potential satisfaction of the individual regarding the content presented.

4. *Users are sensitive about negative information when they read it.*
   It is a general opinion of the users that words that create negative connotations could be avoided from being used in the explanation while being presented to the user.

   - For example, "..extremely low..","..completely hated..", "..avoids places that are heavily disliked by other group members ..", "..unhappy.." created a negative connotation on the users.
   - Additionally, these explanations seemed they were too biased towards the other team members and not supporting the users who had received the explanation.

   Include concluding sentences that "re-assure" the user that they will not be missing out on their favorite places. These pointers will be implemented in the next study.

5. *The structure of the explanation is vital to convey necessary information at the right time to increase potential satisfaction of the user.*
   The general opinion of the participants was to arrange the sentence order in an explanation. This was important to convey a clear message about the sequence items recommended to improve understanding as well as to increase satisfaction.

   The explanation is kept brief by,

- Displaying information describing the situation the user is facing first and then information on the system's transparency describing ratings of users.
- Displaying information on algorithmic transparency in the end.

6. The use of words like "trust", "close friends", "friends" invoke the social bond between group members. Its presence in social explanations can convince the group members better [38]. But, the user comments suggested the presence of such words was not necessary as they already knew their audience. For example, users' comments cited that, *"..if I am already traveling with my close friend, why should the explanation even remind me about it to satisfy me,..."*.

These were some of the analysis and motivations, we found that will help us determine the explanation for the next study.

3.2.5 Summary

This section summarises some do's and do not's that we could potentially consider for determining the content of the explanation. This section also motivates why we could conduct a second study. This would focus on the structure in which the determined content the explanation will be accommodated for every social situation.

- The explanations could avoid the use of "strong words" but choose to convey the same meaning through user-friendly words and phrases. It is noted that this was the same inference about the **choice of words** found in the work of [35].
- Highlighting on the importance of avoiding words or phrases that induce "**negative connotation**" to the whole explanation. The users felt bad and not satisfied when the explanation contained negativity words.
- Further, the use of user's rating preferences as real numbers in the explanation created unnecessary confusion for the user as they had to remember numbers and place names and then understand the justification that the system gave for a particular situation where they could disagree. Hence the **conversion of rating values into categories** of "low, "medium" and "high" could make a good difference in the understanding and, they evaluated these variants of explanations with more positive comments.
- Finally, the explanation could **not just convey information about shortcomings of the recommended sequence** to the user in the pretext of being system transparent. This made the users feel that the recommender system was biased towards them and supported other members. Hence the explanation could convey both positive and negative aspects of the recommended sequence. This could help them understand that the RS was not biased.

3.2.6 RQ 1 - Discussion

RQ 1: How should one formulate the explanation for user satisfaction?

The preliminary study 1, gave following contributions from qualitative analysis of supportive user comments on how the content of the explanation could be presented while trying to express group member's preferences. The contributions are:

1. Reduce the usage of words that express strong emotion.
2. Reduce words with negative connotations.
3. Convert rating values into categorical values.
4. Explain the entire sequence to convey both positive and negative aspects of the recommended sequence.

These results are not supported by statistical values but based on the inferences made from the user comments. This concludes the first preliminary study for determining the content of the explanations.

3.3 Preliminary Study 2 : Structure

The research question addressed in preliminary study two is: *How should one formulate and structure explanations for different relationship strength?* This study has been designed and conducted to understand how to formulate explanations for different relationship strengths and to structure the explanations for the entire recommended sequence. The first part is achieved by providing an explanation to both primary and secondary group types and then analysing user comments. As discussed in literature, primary group are a set of closely related people and secondary group are a set of people who do not know each other on a personal level. Then the latter part is achieved by analysing user comments about the likes and dislikes about the position of blocks of information on algorithmic transparency, names and rating preferences, social situation under discussion. The following sections has discussed the design and materials of the study, procedure that is undertaken by the participants and, results and discussion.

3.3.1 Materials for the study

The explanation we determined for social situation "places I like is not recommended in preferred order" as a result of this preliminary study one is shown below. This is for a group member named *Adam* belonging to a three-member group.

Explanation : "*Dear Adam, you are not recommended some of your favorite places B and J initially in the trip. This is because your travel companions Mary and John have given a low rating to these places. Also, the system considers the lowest as well as the highest rating given by other group members for the place recommended. However, you will visit these places you like later during your trip. We also believe that Mary, who has not rated place B very highly, is compromising in a similar situation like this*".

3.3.2 Study Design

A total of 12 students of TU Delft participated in the study belonging to 4 different nationalities. They were aged between 22 and 27 years and were 40% male and 60% female. This study collects data from the users using a questionnaire that has been created using PHP and HTML programming languages and then the pages were hosted locally. The study is a within-subject study design, where all the participants are exposed to both the group types and they receive an explanation. The participants individually evaluate the explanation in both the scenarios via a questionnaire.

3.3.3 Procedure

The participants gave their consent to participate in the study. Then, the participants were briefed in writing about the research's motive and, basic definition and use of the explanations. For the within-subject study, the participants were asked to assume they were traveling with two of their close friends for scenario 1, and two of their acquaintances whom they have just met for scenario 2 The recommended sequence of POI (E,F,H,D,J,B,G) was shown. Then the participant received an explanation to be evaluated, for both the *scenarios* in the following way:
Scenario 1: Primary group receives an explanation
Scenario 2: Secondary group receives an explanation

Then the participants were asked 2 questions for each of the scenarios listed here. The questions were:

- Tell us what you like most about this explanation given you are traveling with your close friends?
- What would you change in the explanation given you are traveling with very close friends?

The term "close friends" shown for the primary group in the questions was replaced with "acquaintances" for scenarios 2, in both the questions.

3.3.4 Results

The participants subjected to the within-subject study provided us with user comments on both scenarios. Therefore all the 12 participants answered the open-end questions when all of them belonged to each group separately. A qualitative analysis of user comments is done. Some of the comments given by both the groups were focused on the improvement of content and the structure of the explanation. Hence the comments are segregated under the corresponding themes/labels for content and structure. The labels are positive affect, algorithmic transparency, rating preferences, information regarding assurance and order of explanation. Whereas few comments were particularly given expressing the need for privacy, which has been separately discussed under the label "Need for Privacy".

3.3.4.1 Positive Affect

This theme is for discussing the improvements for the *content* and the overall *structure* of the explanation.

Comments: The explanation's sentence phrase "you are not recommended some of your favorite places B and J initially in the trip" was not received positively by participants. They felt it induced negative connotations. One of the users' comments that supported positive connotations were the suggestion *".. hope you look forward to* *your favorite places and have a good time..".* After analyzing the user comments, this issue could be solved by restructuring the phrase like "*place B and J, liked by you, is visited later in the trip*".

Furthermore, the rating preferences focusing on other group member's preferences could make some of the group members feel prejudiced about other group members at the start of the trip. They felt left out of the group. The supporting user comments were *"..since I am not getting the places I like, a positive opening will satisfy me.."*, *"..it feels that you were completely excluded from the decision making or your viewpoints were not considered at all..".* This could be resolved by reminding the user about the favorite places that have been recommended initially on the trip.

Guideline: It can be inferred that the sentences could be phrased on a positive note and could convey information about disagreement situations (arising out of sequence order) after explanation about recommendations that convey information about favorite places that have been recommended.

3.3.4.2 Algorithmic Transparency

This theme is also for discussing the improvements in the content and the overall structure of the explanation. The algorithmic transparency could give a heads-up (idea) of how the entire sequence was computed using one/more of the aggregation strategy stemming from social choice theory. The algorithmic transparency was explained in user-friendly words to avoid technicality. The explanation was *"Also, the system considers the lowest, as well as the highest rating, is given by other group members for the place recommended".*

Comments: This kind of explanation got positive and negative comments from both the group types. The given algorithmic transparency did the work of conveying the necessary working of the system and this conclusion is supported by user comments such as (1) ".. system considers every member's favorite place and preference..", (2) "system tells about how it provides recommendation", (3) "..liked the algorithm's transparency..".

But there were instances where it did not deliver the entire working of the algorithm by not including technical words like "average", "votes" etc. The supportive user comments such as (1) "..need better word choices to understand the unbiased recommendation..", (2) "..should be positively confirmed about the working of the system, so that people know for user that the system is unbiased when we are traveling with strangers..", (3)"..does not mention how it is providing the recommendations like averaging or popularity so that I know system is not biased..", (4)"..should be positive it is unbiased towards all group members..", (5)".. I would like to know what is happening before you tell me about group preferences, otherwise it is confusing.."

Guideline: This issue can be resolved by providing an explanation such as "*the system tends to average out the highest and lowest rating of individual preferences of group members when it is a recommending sequence of places*". It can also be inferred that displaying system transparency about group member's preferences after algorithmic transparency will help the users understand the rating preferences with a context. Additionally, the users preferred explanations that took almost every member's preferences into account.

3.3.4.3 Need for Privacy

The names of group members were the most noticed aspect of the explanation given for both primary and secondary groups. This is because the user's relationship strength with group members differed and names seem to affect their privacy. This theme is to understand if the secondary group members preferred privacy with other group members.

1. **Explanation given to a member of the Primary Group:**
   The explanation that conveyed the users' preference looked like *"This is because your travel companions **Mary and John** have given a low rating"*.

   Comments: The primary group felt the explanation gave them a *"personal aspect"* by displaying the names of the other group members who rated the places they liked low. On the contrary, some participants who received a transparent explanation and assumed they were traveling with close friends gave comments like (1)".. lack of anonymity might create conflicts within groups of friends..", (2)".. names can cause animosity and social discomfort..".

   Guideline: It can be inferred that revealing names were sometimes not preferred amongst the close group of friends.

2. **Explanations given to a group member of the Secondary Group:**
   Explanations given to acquaintances received very few positive comments on the aspect of names. Few participants gave comments close to "like to see the names of group members". On the contrary, as the participants assumed the group members to be acquaintances (who are people with whom they do not share personal relationships) gave comments that were against the disclosure of names of people who rated low. This could be because their own privacy could be compromised when other group members received similar explanation.

   Comments: The comments were "..dislike to see the names of group members who are not close, I prefer to be anonymous..", "..do not name anyone personally..", "..need diplomacy, otherwise this may create awkward situations within the group..", "..more comfortable if travel companion was given instead of names..". Therefore, using names to explain rating preference was considered as a privacy breach when people were with acquaintances.

   Guideline: It can be inferred that anonymity was preferred collectively by secondary group members. This will be evaluated in user study 3.

3.3.4.4 Rating Preferences

The rating preferences were explained using explanations like "*Mary and John have given a low rating to these places*". This is aggregated predictions style of explanation (see Section 2.2.3.2 from Chapter 2). The participants expressed that this explanation gave them insights about the outcome of the sequence and helped them understand the preferences of friends or acquaintances clearly. The categorical values were appreciated for reducing the overhead of remembering numbers.

Comments: The supportive user comments were (1) ".. clearly specifies the outcome of the sequence and tell the reason why", (2) "..explanation is analyzing each user's interest..", (3) "..like to see how much people rated for each place..", (4) "..interesting to know what others are preferring while traveling..", (5) "..like to see choices of my close friends..", (6) "..like the way low-medium-high is used to explain the rating directly and not beating the bush..", (7) ".. interesting to know preferences of acquaintances ..".

The participants expressed their *dislike and curiosity* through the user comments such as (1) "..choice of rating could be due to time constraint/cultural difference..", (2) "..makes me think John and Mary who are my friends did not like my choice of places at all, this makes me sad..", (3) "..not liking that the explanation is solely based on recommendations, even though it is nice to know the preference of acquaintances..".

Guideline: We can infer that explaining preferences by citing the person responsible for the lowest and highest rating could not be avoided much. Therefore it is advisable to adapt the "feedback sandwich" method [11]. Hence rating preferences that induce negative connotations could be surrounded by information about favorite places recommended and positive assurance.

3.3.4.5 Information regarding Assurance

Assurance is about giving confidence to the users about their favorite places is/will be visited. For instance, information of assurance is actually an outcome of the sequence but is not inferred easily from the recommended sequence's order. The participants expected *more insight and assurance* that they will not miss out on their favorite places. This information is presented as *"However, you will visit the places you like later during your trip. We also believe that Mary who has not rated place B highly is compromising in a similar situation like this"*.

Comments: The supporting user comments were "..like the fact that places you are missing will be visited later..", "..nice to know other group members are compromising..", "..good that the system informs me about this, it makes me think about others..". The negative comments such as "..hard to believe that some compromise..", "..this information is sometimes unnecessary..". This also poses a challenge to the introduction of the concept of explaining about compromise.

Guideline: It has been inferred that, informing about this kind of assuring information can be kept to the end of the explanation to provide a holistic approach to the structure. Thus explanations that only inform about member's preferences can be supported with information on assurance.

3.3.4.6 Information Order in explanations

To determine the potential structure of the explanation irrespective of the group type, each of the social situations needed an explanation in order to explain the full sequence order of recommendations.

Comments: The user comments which inspired this idea were (1) "..I would like to see when would the places I love the most is coming up next and then only followed by why it was not recommended due to group's preferences..", (2) "..later it feels a bit negative with indirectly saying that the group needs are bigger than yours..", (3) "..explanation is biased and does not consider my preferences..", (4) ".. I would like to know what is happening before you tell me about group preferences, otherwise it is confusing..", (5) ".. end on a positive note as the explanations should convey a message yet covey the hard truths about the group's rating..".

Guideline: One of the main keys to increasing satisfaction could lie in the way the information is conveyed in a convincing manner. The facts about the disagreement situations like "places I like is never recommended" and "places I like is not recommended in the preferred order" could be safely cushioned by a positive situation like "place I love the most is recommended" and algorithmic transparency in the start and then information of assurance and compromises made by others in the end. This is made to give a holistic look at the explanation given to the entire sequence. This way of arranging any information is called *sandwich model* [11].

3.3.5 Summary

In this study, some inferences were made on the potential structure, the content of the explanation and the privacy requested by the participants.

- **Content:**
  - The *disclosure of names* in the transparent explanation amongst friends helped the users feel a personal aspect. But participants suggested that the presence of names may create social discomfort or targeted animosity even amongst friends
  - The aspect of names created a privacy breach for acquaintances and they requested anonymity. This was to maintain diplomacy and avoid awkward situations. The phrases such as "some members", "few members", "travel companions", the number of members were suggested in the place of names. From this, it is shown that there was a need for privacy
- **Structure:**
  - The explanations about the disagreement with the order of items could be phrased on a positive note. This information can be conveyed after informing the member about their favorite places that could be recommended at the start of the sequence.
  - Since conveying the information about the lowest and highest preferences of group members could not be avoided, a "feedback sandwich" model [11] could be adopted for explaining the whole sequence. That is the explanation about members receiving their favorite places could be informed prior to the information about disagreement situation.
  - The *system transparency* about member's preferences on items could be understood well after the users are exposed to **algorithmic transparency**.
  - Assuring information about other members facing similar kinds of disagreement with the order of the sequence will help in providing a holistic approach to the explanation. This conveys the information about the recommended sequence that was not easily interpretable.

3.3.6 RQ 2 - Discussion

The explanations were presented to both primary and secondary groups hence depicting two scenarios that were presented in the within-subject study. After evaluation of the explanations from both the scenarios separately, the participants gave their comments. We have performed qualitative analysis on the user comments from both the scenarios and made inferences (refer Section 3.3.5). The inferences made for the explanation's structure for a social situation helped us to structure the explanations for other social situations as well and make an explanation for the entire sequence. Due to the need for privacy expressed by the participants of this preliminary study 2, an explanation with preserved privacy was designed by reproducing the explanation and removing names.

3.4 Conclusion

We have learned about the content and order of the information in the explanation. We also learned that privacy was important. Therefore we will have two versions of explanations. The explanation with entire information that has names and rating preferences will be henceforth called "transparent explanation" and the version of the explanation with no names will be called "privacy-preserved explanation". In the next chapter we will evaluate the explanation with the following structure.

1. Algorithmic Transparency
2. Positive Reassuring Explanation about "Places you like is recommended"
3. Repair Related Explanation about "Places you like is never recommended at all"
4. Repair Related Explanation about "Places you like is not recommended in the preferred order"
5. Repair Related Explanation about "Places you dislike is recommended"
6. Positive Assuring explanation about "People who also had to compromise".

Points 3, 4 and 5 are arranged in the decreasing order of situations that may cause negative connotations. Here point (2) reminds them about the highly-rated places the member preferred. Then point (6) reminds the member about the sacrifices or compromises the other group member also has to go through when they are accommodating others' preferences presented to them using points (2-5).

As seen in the literature (Chapter 2), the explanation that focuses on the difference in opinion among the members of the group by reminding them about the preferences of other group members for the places recommended [35]. Also, explanations that convey information on situations where all the group members agree with each other is called reassuring explanations [35]. This thesis has continued using this terminology.

Chapter 4

Evaluating Explanations for different relationship strengths

4.1 Introduction

Motivation: In Chapter 3, the content, and structure of transparent explanations were determined. This explanation was replicated and names were hidden to make it privacy preserved. The results of user study two, suggest that both group types requested privacy when participants were given transparent explanations. This *motivates* us to evaluate whether secondary group type (acquaintances) preferred privacy preserved explanation in comparison to transparent explanation. Similarly, there is a need to investigate whether primary group type (friends/couples) preferred transparency to privacy.

Therefore the research question that is focused through this final user study is: *Do different explanation types for different relationship strength influence user satisfaction?*. The results of this study will help us determine whether privacy was preferred only when people traveled with acquaintances as suggested in the literature and preliminary study 2; or on the contrary, whether this was the case with friends as well.

For this study, participants will evaluate each of the explanation types given to each group type. This means each participant will be subjected to four scenarios each (within-subject study). They evaluate the explanations through a questionnaire where they rate their satisfaction with the given explanation for the recommended sequence on a scale of 5 (1 being highly dissatisfied and 5 being highly satisfied). They also give written user comments expressing their likes and dislikes with the given explanation when they belong to each group type. The numerical analysis will be made on the satisfaction value and qualitative analysis will be made on the opinions of participants to further investigate which explanation satisfied them the most when they belonged to a particular group type.

4.2 Materials for the study

The version of explanations presented to the user in User study 3 are as follows.

1. **Transparent Explanation:** Dear Adam, the system tends to average out the highest and lowest rating of individual preferences of group members when it is

a recommending sequence of places. It recommends places that make sure that no-one is unhappy. We see that the first few highly rated places (D, E, F, and H) in the recommended sequence are according to your interests. But there are some situations where you may disagree with other group members. We note that place C is rated highly by you, but it has not been included in the sequence. This is because **place C is rated low by Mary and John**. Additionally, we see that place B and J, liked by you, is visited later in the trip. This is due to the fact that this **place B is rated low by John and place J is rated moderately by Mary**. Finally, place G is rated low or moderately by you and other members in the group hence it is recommended later in the sequence. Although you are recommended the places you don't like, you get to visit places you love first. Furthermore, almost all group members, are compromising in similar situations of disagreement like these to satisfy the group. Hence this method is the best way to maximize the group's satisfaction.

2. **Privacy Preserved Explanation:** Dear Adam, the system tends to average out the highest and lowest rating of individual preferences of group members when it is a recommending sequence of places. It recommends places that make sure that no-one is unhappy. We see that the first few highly rated places (D, E, F, and H) in the recommended sequence are according to your interests. But there are some situations where you may disagree with other group members. We note that place C is rated highly by you, but it has not been included in the sequence. This is because place C is **liked by only a handful of members**. Additionally, we see that place B and J, liked by you, is visited later in the trip. This is due to the fact that place B and J are **rated moderately only by a handful of people**. Finally, place G is rated low or moderately by you or other members in the group hence it is recommended later in the sequence. Although you are recommended the places you don't like, you get to visit places you love first. Furthermore, almost all group members, are compromising in similar situations of disagreement like these to satisfy the group. Hence this method is the best way to maximize the group's satisfaction.

Both of these explanations can be made for every other group member (here, Mary and John) separately. We have shown the explanation designed only for Adam for clarity. These explanations are shown above and are proposed to the participants (n = 25) of the user study 3. These participants will evaluate each explanation when they belong to each group of different relationship strength separately.

4.3 Study Design

A total of 25 participants aged between 22 and 55 years of whom 54% were male and 46% female. People from 10 different countries took part in this experiment (Chinese, Costa Rican, Dutch, India, Indonesia, Nigeria, Portuguese, Russia, Singapore, and Germany). Therefore the data is not biased.

Each of the participants (n = 25) was subjected to a within-subject study. The Figure 4.1 displays four scenarios which explain how the study took place for each

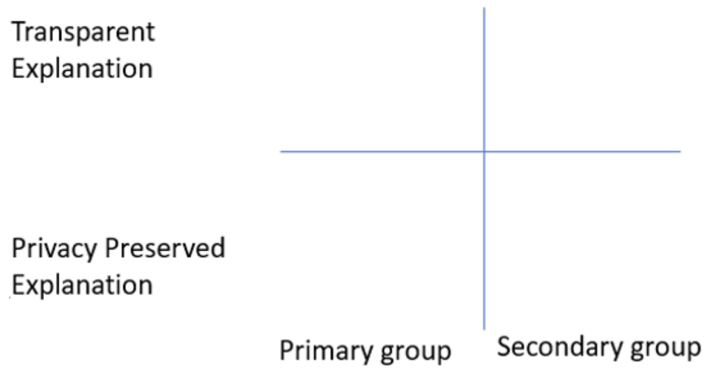

Figure 4.1: Four quadrants depicting the scenarios that were presented to every participant

participant. Every participant was subjected to all four scenarios. The participants received either transparent or privacy-preserving explanation when they were subjected to one group type at a time. The scenarios were:

- *Scenario 1:* Primary group receives Transparent Explanation
- *Scenario 2:* Primary group receives Privacy-Preserving Explanation
- *Scenario 3:* Secondary group receives Transparent Explanation
- *Scenario 4:* Secondary group receives Privacy-Preserving Explanation

All four scenarios had a questionnaire. Therefore each participant answered the three questions (refer Section 4.4) in the questionnaire in each of the four scenarios. The questionnaire was conducted in locally hosted HTML and PHP web-pages. The order in which the scenarios were presented to the users were changed frequently, to reduce similar "carry-over effect" from one experimental condition to other. The answers to the questions of this user study are used to test the hypothesis and provide further analysis in this experiment.

4.4 Procedure

The participants were requested their consent to take part in the experiment. Then the research's goal and motive, the basic definition of explanations were briefed at the start of the study. Next, they were given a brief introduction about how explanations were proposed for the sequence of recommendations given to them. Then a picture of friends is shown for scenario 1 and 2 and picture of recent recruits meeting for the first time on a coffee break is shown for scenario 3 and 4. Then the picture's presence is explained to the participants. This was only to remind the participants they belonged to the primary group for the first two scenarios and the secondary group for the last two scenarios. Each participant was asked to assume that they were with two of their close friends for scenario 1 and 2 and with two of their acquaintances with scenario 3 and 4. Then the recommended sequence is shown which is from the same algorithm as mentioned in Section 3.1. Then this sequence was explained with either transparency and privacy preserved explanation as shown in Section 4.2. The participants were asked three questions for every scenario through which they evaluate the explanations based on their satisfaction. The **questions** were:

1. Are you satisfied with this explanation?
2. Tell us what you like most about this explanation given you are traveling with your acquaintances?
3. What would you change in the explanation given you are traveling with very acquaintances?

The word "acquaintances" was replaced with "close friends" when participants were subjected to the primary group. The answer to the first question was chosen from a 5 point Likert scale (1 being extremely dissatisfied to 5 being extremely satisfied). Statistical analysis is done on this value. The satisfaction degree measured by question 1 is based on how the users react about the *presence or absence of privacy (names)* in the given explanation for a particular group type. The written opinion given by the users for the next two questions (2,3) is used for qualitative analysis to further support the conclusions that result from testing the hypothesis mentioned in Section 4.6.

4.5 Variables

The independent variable would be the transparent and privacy preserved explanation and the primary and secondary groups. The satisfaction felt by the groups is measured for every explanation for every group and hence this is the dependent variable.

4.6 Hypothesis

The two hypotheses designed to answer the research question are:

1. **H1:** The primary group prefers Transparent Explanations rather than Privacy preserved explanations as it will satisfy them better.
2. **H2:** The secondary group prefers Privacy-Preserving explanations rather than Transparent explanations as it will satisfy them better.

Hypothesis 1 and 2 can be tested via statistical test. For primary group (H1), there is a check for significant difference between the mean values of the satisfaction degree of both explanation types (transparent and privacy-preserved) given by the participants. A similar observation is done for the secondary group (H2).

4.7 Results

The user studies have helped us measure the user satisfaction of the participants via satisfaction values given by them for each explanation type on a 5 point Likert scale. This is done for all four scenarios. The results for each hypothesis is discussed here.

For clarity's sake, the same 25 participants took part in all the four scenarios. Hence 25 participants were present in the primary group and they were subjected with explanations types separately in scenarios 1 and 2. This was the same case for the secondary group in scenario 3 and 4.

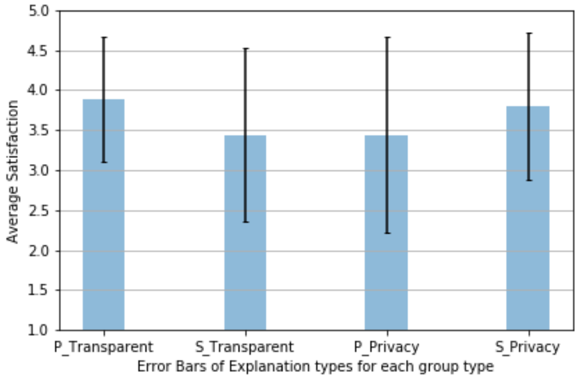

Figure 4.2: Average Satisfaction received for each Explanation type by both Primary and Secondary group types, Whisker plot of almost 1 SD.

From Figure 4.2, S_Transparent and S_Privacy is the average satisfaction value of transparent and privacy-preserving explanation given by secondary group respectively. Similarly, P_Transparent and P_Privacy is the average satisfaction value of transparent and privacy-preserving explanation given by primary group respectively.

H1: Primary group prefers Transparent Explanations rather than Privacy preserved explanations as it will satisfy them better.

The mean and standard deviation of the satisfaction values for both explanation type is calculated for the primary group. The transparent explanation has μ = 3.88 ,σ=0.781 and privacy-preserving explanation has μ= 3.44, σ=1.22 (refer Figure 4.1). The satisfaction values for both the explanation types were not normally distributed. Wilcoxon signed-pairwise test was conducted on the satisfaction values of transparent and privacy-preserving explanation given by **primary** group. The p-value is 0.0954 and the alpha value is 0.05, the reported statistic is 56.0. The p-value is greater than alpha value, therefore we fail to reject the null hypothesis of equal means. There is **insignificant statistical difference** between the average satisfaction values. We see that the primary group does not have a preference over the explanation they would like. Figure 4.2 show standard deviation of almost 1 for both explanation types.

Therefore we are not able to support H1.

H2: Secondary group prefers Privacy preserved explanations rather than Transparent explanations as it will increase their satisfaction better.

The mean and standard deviation of the satisfaction values for both explanation type is calculated for the secondary group. The transparent explanation has μ = 3.44

, σ=1.08 whereas privacy-preserving explanation has μ = 3.80 , σ=0.912. The satisfaction values for both the explanation types were not normally distributed. Wilcoxon signed-pairwise test on the satisfaction values obtained for the transparent and privacy-preserving explanation for **secondary** group. The p-value is 0.135, the alpha value of 0.05 and, the reported statistic is 74.5. The p-value is greater than alpha value, therefore we fail to reject the null hypothesis of equal means. There is **insignificant statistical difference** between the average satisfaction values. We see that the secondary group does not have a preference over the explanation they would like. Figure 4.2 shows high variance in the standard deviation of almost 1 for both explanation types.

Therefore we are not able to support H2.

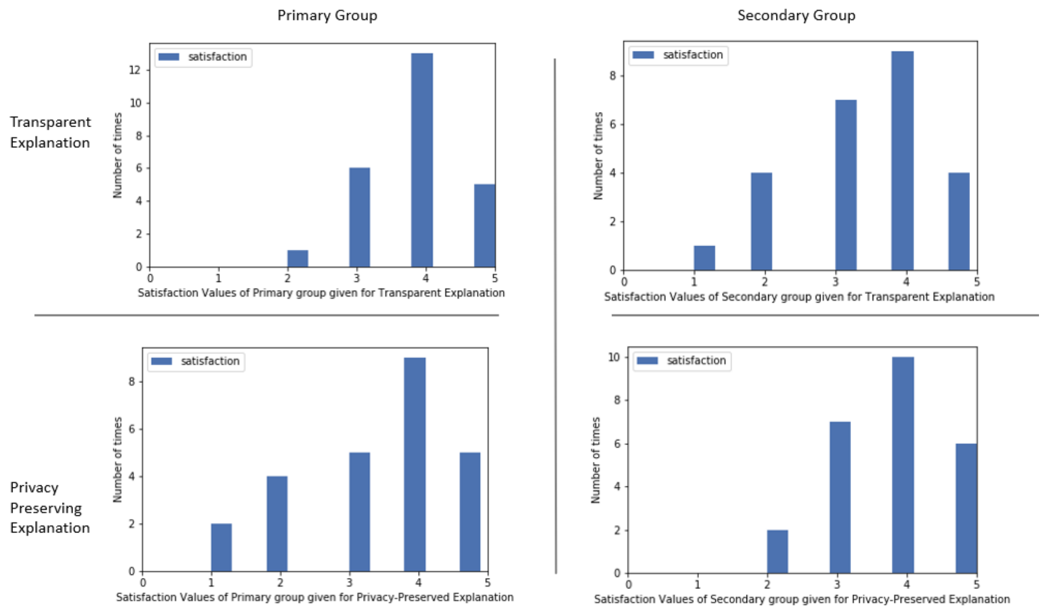

Figure 4.3: Satisfaction Values expressed by Primary and Secondary groups for Transparent and Privacy-Preserved Explanations

Overall Satisfaction: The satisfaction values expressed by both group types for both explanation types were voted a 4 or a 5 by most of the participants (see Figure 4.3). The overall average satisfaction of the explanation types across both group types presented in the main experiment was scored 3.6 on the Likert scale of 5 (see Table 4.1). This indicates that explanations on a whole across all scenarios had satisfied the participants due to the word choice alias content and structure in which it was presented. Hence study 1 and 2's contributions to design satisfactory explanations are justified.

Comments for overall satisfaction: The comments such as ".. explanation was detailed..", "..offers enough explanation to justify the travel plan..", ".. I like the basic reasoning of the system and how the explanation tries to ensure how the recommendation related to my own preferences..", ".. like to see that I would get the places I love first..", "..clear explanation of logic why recommendation system is suggesting certain result..", ".. explanation makes me think about all others preferences..", ".. explanation gives justification for the choice of recommendation..", ".. like the way it tends to showcase how groups compromise.." also add enough evidence that the explanation, on the whole, is quite satisfactory.

4.8 RQ 3 - Discussion

RQ 3: Do different explanation types for different relationship strength influence user satisfaction?

Sample size: We had 25 participants. Therefore we conducted a power analysis test (Post hoc analysis - Wilcoxon signed ranked test). Power is the probability of finding an effect (i.e., getting a p-value of < .05), given that there is an effect. The effect size considered was Cohen's d. This is done for the explanations given for the primary group and the same was repeated for the secondary group. With α=0.05 and two-tailed test, the hypothesis H1 had an effect size of only -0.411 with 48.5% power that the alternative hypothesis of unequal means would have been accepted. Whereas the hypothesis H2 had an effect size of only 0.355 with 38% power that the alternative hypothesis of unequal means would have been accepted. This suggests that if the sample size was larger; the difference between the mean that was found or an extreme value occurring had a lesser effect on the hypothesis. Therefore the results were not significant for H1 (test statistic = 56 , p = 0.0954, d = -0.411) and, for H2 (test statistic = 74.5 , p = 0.135, d = 0.355).

Table 4.1: Statistical Results - There was no significant difference between satisfaction's mean values for given for the explanations calculated across each group

Statistical Results	Primary Group	Secondary Group	Overall
Transparent	μ = 3.88 ,σ=0.781	μ = 3.44 , σ=1.08	μ = 3.66,σ= 0.9305
Privacy Preserved	μ= 3.44, σ=1.22	μ = 3.80 , σ=0.912	μ = 3.62 ,σ= 1.066
Overall	μ=3.66 , σ=1.0005	μ =3.62 , σ=1.0	—-

Trend, Mean, SD: When we observe table 4.1, it is interesting to note the slight difference in means of both explanation types for each group type (even though it is insignificant). Therefore, the trends are in the direction that has been predicted in the hypothesis. Also, we notice the mean is lower for the transparent explanation given to secondary groups and when privacy preserved explanation was given to primary groups. This visible dip in the mean suggests it may due to the fact that secondary group feels a privacy breach while the primary group feels the restriction of not knowing information about the likes and dislikes of their friends.

Due to an insignificant difference in average satisfaction, we are unable to conclude, whether the individuals in the primary group are satisfied with transparent explanations (with names of group members and respective preferences) or privacy-preserving explanations better. This is the same case for the secondary group.

There is a **high variance in standard deviation** (around 1 SD) in all the explanation types given to group types. The user comments suggest, that *each individual has different opinions* on the explanation type given to them under each group type. This could be due to different ways in which an individual could handle conflict within the group due to the disagreement with the order of the sequence recommended. Furthermore, it could be due to the reason that the explanations collectively explain the three different situations in which conflict can occur. In the posthoc analysis, analyzing different ways of dealing with conflict can help us understand the reasons high standard deviation.

4.8.1 Post-hoc analysis

Motivation: Before, we saw that satisfaction measured for different types of explanations had a high standard deviation. To address the high variance in standard deviation and difference in opinion displayed by the user comments, we decided to investigate the reason behind it. For all the four scenarios, the user comments of a certain set of participants showed a pattern that was different from the pattern displayed by other sets of participants. The pattern is about their opinions about whether they would like privacy or transparency when they traveled with primary (friends) and secondary (acquaintances) group. To understand this, we analyzed the user comments and found that participants had individual differences. The way each individual deals with conflicting situations, when they traveled with friends or acquaintances, seem to vary. For clarity's sake, conflicting situations occurred due to the individual's disagreement with the recommended sequence order of POI. Since, we detected that personality could play a role in how group types reacted to the explanation type give, this could be considered as *co-variate*. Personality as an independent variable could be affecting the satisfaction values we have measured for each group type.

Co-variate: All the participants were requested to take "Thomas-Kilmann Conflict Mode Instrument (TKI) personality assessment" test [28] at the end of the user study 3. The purpose of the TKI test is to help an individual understand how they behave when confronted with a conflict situation [19]. This test also controls "social desirability" which is a type of response where a person answers questions so that the answers are looking favorable to others [19]. The test helps people manage conflicting situations in a positive manner. Hence, the test will help us determine the personality profile of each user.

TKI test is made of 30 questions. The participant should choose one out of two options displayed for each question. The answers to the test are evaluated and raw scores and percentile scores are assigned to every *conflict mode style*. The raw scores ranged from 0 until 12 for every conflict-mode style [19]. The percentile scores were 25, 50 or 75 were assigned to each raw score computed based on the "TKI raw score and percentile for international sample" table found in [19].

Conflict modes: The authors of [28] have identified five "conflict-handling modes" that people use while dealing with conflict. They are (1) *competing*, (2) *col-*

laborating, (3) *compromising*, (4) *avoiding*, and (5) *accommodating*. The five different conflict-modes are shown in Figure 4.4 with cooperativeness (satisfying other's preferences) in the x-axis and assertiveness (satisfying one's own preference) on the y-axis [19].

The Competing mode is assertive of their own preferences and very less cooperative with other group members' preferences. The collaborating mode would like to find a definitive middle ground and arrive at a "win-win" solution by satisfying all group member's preferences Avoiding conflict mode usually "side-steps" the conflicting situation without trying to satisfy their own concerns or of the other members. Accommodating conflict mode self sacrifice their own preference for the sake of other group members. The compromising mode gives up comparatively more than competing personality, but less than accommodating. They also do not sidestep as much as avoiding and do not look at everyone's preferences very clearly like collaborating [43]. Therefore, compromising mode would only partially satisfy their own preference as well as other group member's preferences. The definitions given in this paragraph is from [43].

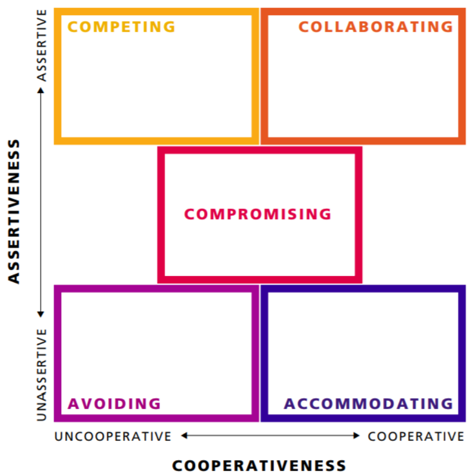

Figure 4.4: Conflict Mode Style - Thomas-Kilmann Personality Assessment test [43]

Metric: An individual will have a combination of all the five conflict modes and cannot be deemed to have just one of the 5 modes as his/her character [43]. Some of them use some modes readily or some have a "clear favorite". The way a person deals with conflict can be based on the situation and the mode they have been using for a very long time [43]. Therefore, we decided to compute the predominant behavior of an individual by assigning the most frequently used conflict modes to the individual and then segregate their comments according to the conflict mode to find a pattern. All the modes that have received a percentile score of above 75 percentile have been chosen for each participant.

Note that, On assigning all dominant personalities that has been determined for each participant, there was an unequal distribution of personality types in this study. For example, user 1 would have compromising, avoiding and accommodating, whereas user 2 has a clear favorite of avoiding, user 3 has a combination of collaborating and avoiding and so on. This was not the case in [37]. Quijano et al., [37] in their work, assigned twenty users for each type of personality (very selfish, selfish, tolerant, cooperative, very cooperative). Then they made groups with a predominant personality type assigned to each group member to conduct experiments to observe how each personality behaved.

4.8.1.1 Analysis of Conflict modes with the support of user comments

For each conflict mode, a summary of *general* preferences and *specific* preferences expressed when the particular conflict mode traveled with each group type is discussed. The specific preferences denote preferences according to group type (relationship strength).

1. **Avoiding conflict mode**
   - **General:** This conflict mode (a) did not care about the order of the sequence, (b) did not see a necessity for an explanation regarding situations of disagreement, (c) were ready to change their preferences about a place, to avoid conflict.
   - **By Group-Type:**
     - This conflict mode has rated transparent explanations *comparatively less* when they were to travel with *acquaintances* than with friends.
     - Even though this mode disliked transparency, it was tolerated with friends. This is because they were ready for a discussion to make a better plan that is pleasing to all members.
     - With acquaintances, transparent explanations made them experience a breach of privacy as there was no anonymity and created a situation of unplanned discussions, which they were not interested in.
     - Privacy-preserving explanations were appreciated as they provided expected privacy.
2. **Competing conflict mode**
   - **General:** This conflict mode (a) preferred usage of affirmative word choice and, (b) were confident that their friends would not disagree with their preferences.
   - **By Group-Type:**
     - Transparent explanations helped this conflict mode to know their group members better. These explanations gave them enough information to either convince others to agree to their plan or discuss a better plan.
     - Privacy-Preserving explanation hindered the chance to get to know other group members.
     - They seem to prefer transparent explanations when they had to travel with either of the group types.
3. **Collaborating conflict mode**

- **General** This conflict mode likes that the explanation logically explained how places were recommended so that no-one was unhappy (average without misery strategy - refer Section 2.1.1.3).
- **By Group-Type:**
  - They seem to prefer transparent explanations when they had to travel with either of the group types.
  - The presence of names and ratings were important with friends and acquaintances to clearly understand whether every group member was satisfied as they were.

4. **Accommodating conflict mode**

   - **General:** This conflict mode feels (a) the explanation of the logic (algorithm) behind the recommendation makes this personality comfortable. This is because the very cooperative nature of this mode, wants to make sure, everyone is satisfied in the group by receiving their preferences, (b) the information about everybody in the group compromising with their favourite places at some point of the trip (order of the sequence) is appreciated, (c) the explanation of the entire itinerary is helpful, (d) likes that the explanation for places which are not recommended in preferred order is given
   - **By Group-Type:**
     - They seem to *follow the trend as mentioned in the hypothesis*.
     - This mode seems to prefer transparency with friends so that they can comply with the preferences of others more easily.
     - They appreciate privacy with acquaintances, as the explanation showcases that the group compromises without revealing identity.

5. **Compromising conflict mode**

   - **By Group Type:**
     - They seem to prefer transparent explanations.
     - To initiate a discussion after receiving explanations this mode prefers transparency to privacy.
     - With acquaintances, this mode is ready to accept the order of the recommendation as they have also got their favorite places and not the only ones compromising in the group.

The limitations are that the observed results cannot be concluded as a concrete conclusion from the experiment due to the less number of participants and an insignificant difference in the average satisfaction of the explanations across groups.

4.9 Limitations

1. **Restricted Explanation Length:** In our study explanations were limited to groups with only three users.

- **Solution to reduce length:** The transparent explanation is designed in such a way that there was a reduced explosion in length of the explanation when there could be more than three users. The explanation can begin to look messy with lots of names of people, places and corresponding ratings and hence becomes long to read and comprehend in the end if an adjustment is not made. The **adjustment** is that the explanations will have only *the lowest score* of a recommended item with the responsible member for that score.

  For example, we suggest "place B is rated low by John and place J is rated moderately by Mary." *instead of* "place B is rated low by John and moderately by Mary and place J is rated moderately by Mary and high by John". Additionally, if two or more places are rated the same then it is given as "place C is rated low by Mary and John" *rather than* stating "place C is rated low by Mary and place C is rated low by John". This was mainly to reduce the length of the explanation even when conveying the reason for the low ratings for the item in the sequence.

- Even though the length of the explanation is controlled by mentioning the ratings that are lowest; it will still create a **overhead** when the number of users or the number of places increases in size.

- **Satisfaction** of the Explanation is validated only for group size of three users. The limited group size may also have *amplified* the effect of relationship strength for groups. Furthermore the satisfaction value is given to the question "are you satisfied with the explanation when you are traveling with friends/acquaintances?" could have been due to the satisfaction with the given explanation as a whole or based on how satisfied a participant was with the given explanation when he/she belonged to a particular group type.

2. **Synthetic Ratings:**

   - The experiment lacks **ecological validity** because we used artificial POIs and synthetic ratings. In addition, we simulated the composition of groups.

   - We conducted a controlled experiment with artificial places of interest in order to avoid distraction with nuances of place types. Synthetic ratings were used, instead of asking users to rate the places based on the type of places and their real interest with the given place. A more realistic way of rating places can affect the way people rated places when they traveled with friends and acquaintances.

3. In addition to explaining a member's dislikes, the explanation could have achieved better satisfaction value by expressing the places that other group members like as well.

4. The privacy preserved explanations could have displayed the percentage of people disliking a place instead of using phrases like "a handful of people". This could give a better understanding of the number of people favoring a place.

5. The **chosen social choice strategy** algorithm may affect the satisfaction of the users more than the explanation style. This might explain the similar results for transparency and privacy-preserving explanations.

6. User receives explanations that consist only of the four probable situations. Social situations that are not discussed here in this thesis may be suggested by the user.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

This work has presented transparent and privacy-preserving explanations for groups and evaluated them by measuring the satisfaction value given by the participants when they belonged to groups of different relationship strengths. We have conducted three user studies separately.

For research question 1, preliminary study 1 was conducted for investigating the content of the explanations. This study was conducted with 5 participants to investigate the content that satisfied the users. This study was exploratory. Different versions of explanations expressing the aggregated predictions were presented to the user across four social situations. They gave their opinions expressing their likes and dislikes of each version. The results of this study were to (1) reduce the usage of words that express strong emotions, (2) reduce words with negative connotations, (3) convert numerical rating values into categorical values to reduce overhead, (4) convey both negative and positive aspects of the recommended sequence while explaining the sequence.

For research question 2, preliminary study 2 was conducted to investigate the structure in which the explanations could be presented and also investigate if the group types expressed the need for privacy. This study was conducted with 12 participants and is exploratory. The user comments were analyzed to improve the content, identify a structure to present the content and understand the need for privacy. The results of this study for the structure, content, and privacy were (1) disclosure of names amongst friends could create social discomfort even though it gave a personalized effect to each group member, (2) aspect of names created privacy breach amongst acquaintances, (3) explanations about situations that created disagreement with the order of the recommended sequence could be explained on a positive note, (4) a "feedback sandwich" model could be avoided for explaining the whole sequence, (5) system transparency could be given after algorithmic transparency, (6) assuring information about other group members facing similar issue could help in providing a holistic approach to the explanation.

For research question 3, user study 3 finally evaluated whether privacy or transparency was preferred by primary and secondary group type. This was done by measuring satisfaction value and comparing the average satisfaction values given by the group type for each of the explanation separately. There were 25 participants who took part in all the four scenarios that were a combination of both the group types and explanation types. The overall satisfaction of the explanation types is 3.6 on a scale of 5. This shows that the explanations were quite satisfactory. There was no significant difference between the average was quite high at 1 SD. The posthoc analysis revealed that individual difference present for each user could be one of the potential reasons for high standard deviation. The personality profile of each participant was calculated using the TKI personality assessment test. After a detailed analysis of user comments under different conflict mode style, we have some inferences that suggest how the individual difference made each person react in a different manner when they traveled with friends and acquaintances.

5.2 Future Work

We recommend future recommendations such as (1) Ecological Validity, (2) Group Size, (3) Social Strategy, (4) Generalizability with the possible situations that can occur due to the order of the recommended sequence and, (5) Privacy.

1. **Ecological Validity:** The experiment could have been conducted with a real group giving real ratings to places with real names. Instead of asking a participant to imagine that they belong to a group of acquaintances or friends we could form groups with different compositions. This could invoke a better understanding if relationship strength really affected the satisfaction value of the explanation given for the recommended sequence. Stronger results might have been found but requires further study. It is possible that relationship strength's effect can decrease when the group size is larger. A combination of different relationship strengths can be present. For example, an acquaintance could decide to travel with a couple, or it is possible a family who is assumed to have a close setting, can differ in relationship strength in a real setting. It is also possible that the rating preferences may change in a real group setting.

2. **Length of Explanation:** The length of the explanation designed for the whole sequence increases as the number of group members increases in the number and the number of places recommended increases in length. This issue was addressed in this thesis to an extent by including only the lowest rating and the person responsible for the same. Further, this can be solved by grouping the places assigned to each social situation under a common name. For example, if we are using place names like Louvre, Musee d'Orsay can be grouped under a common name called "museum" to reduce length. Another solution is to present the number of people who opt for a particular place in terms of percentage to preserve privacy yet give the users a fair idea of the majority of the group's preferences. Also, we could develop visual explanations to solve conflict while explaining sequences for the same.

3. **Social Strategy:** We have created explanations based on the combination of three social choice strategies as discussed in preliminary study 1 and evaluated the satisfaction of the explanations for different relationship strengths. A study

can be conducted where aggregated-predictions style explanations can be designed for each of the social choice strategy mentioned in Section 2.1.1.3 and then evaluated for different relationship types. The same can be done for hybrid social choice strategies as well. Then we can understand if the aggregation strategy chosen affected our final results.

4. **Generalizability:** Instead of the "feedback sandwich" model, we could use other models to make a better explanation for the gap on how to make a good explanation. Additionally, the experiments can be conducted with real groups and real ratings instead of synthetic ratings and artificial place names.

5. **Privacy:** It is possible that users of the recommender system feel a privacy breach other than just names that the explanations reveal. For example, explanations that involve demographic information either directly by stating such information in the explanation or segregate the group demographically and giving them an explanation [3] can affect the user's privacy.

Bibliography

[1] Explanation generation using simplenlg. https://github.com/Nivedita-256/Final_code_master_thesis.

[2] Simplenlg github. `https://github.com/simplenlg/simplenlg`.

[3] Liliana Ardissono, Anna Goy, Giovanna Petrone, Marino Segnan, and Pietro Torasso. Intrigue: personalized recommendation of tourist attractions for desktop and hand held devices. *Applied artificial intelligence*, 17(8-9):687–714, 2003.

[4] Kenneth J Arrow. A difficulty in the concept of social welfare. *Journal of political economy*, 58(4):328–346, 1950.

[5] Mustafa Bilgic and Raymond J Mooney. Explaining recommendations: Satisfaction vs. promotion. In *Beyond Personalization Workshop, IUI*, volume 5, page 153, 2005.

[6] Henriette Cramer, V Evers, MV Someren, S Ramlal, Lloyd Rutledge, Natalia Stash, L Aroyo, and B Wielinga. The effects of transparency on perceived and actual competence of a content-based recommender. In *Semantic Web User Interaction workshop at CHI, Florence, Italy*, 2008.

[7] Henriette Cramer, Vanessa Evers, Satyan Ramlal, Maarten Van Someren, Lloyd Rutledge, Natalia Stash, Lora Aroyo, and Bob Wielinga. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction*, 18(5):455, 2008.

[8] Berardina De Carolis, Nicole Novielli, Vito Leonardo Plantamura, and Enrica Gentile. Generating comparative descriptions of places of interest in the tourism domain. In *Proceedings of the third ACM conference on Recommender systems*, pages 277–280. ACM, 2009.

[9] Amra Delic, Judith Masthoff, Julia Neidhardt, and Hannes Werthner. How to use social relationships in group recommenders: Empirical evidence. In *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*, pages 121–129. ACM, 2018.

[10] Amra Delic, Julia Neidhardt, Thuy Ngoc Nguyen, and Francesco Ricci. An observational user study for group recommender systems in the tourism domain. *Information Technology & Tourism*, pages 1–30, 2018.

[11] Anne Dohrenwend. Serving up the feedback sandwich. *Family practice management*, 9(10):43, 2002.

[12] Alexander Felfernig, Ludovico Boratto, Martin Stettinger, and Marko Tkalčič. Explanations for groups. In *Group Recommender Systems*, pages 105–126. Springer, 2018.

[13] Alexander Felfernig, Ludovico Boratto, Martin Stettinger, and Marko Tkalčič. *Group Recommender Systems: An Introduction*. Springer, 2018.

[14] Alexander Felfernig, Sergiu Gordea, Dietmar Jannach, Erich Teppan, and Markus Zanker. A short survey of recommendation technologies in travel and tourism. *OEGAI journal*, 25(7):17–22, 2007.

[15] Stettinger M. Tkalčič M. Felfernig A., Boratto L. *Algorithms for Group Recommendation. In: Group Recommender Systems*. SpringerBriefs in Electrical and Computer Engineering. Springer, Charm, 2018.

[16] Mike Gartrell, Xinyu Xing, Qin Lv, Aaron Beach, Richard Han, Shivakant Mishra, and Karim Seada. Enhancing group recommendation by incorporating social relationship interactions. In *Proceedings of the 16th ACM international conference on Supporting group work*, pages 97–106. ACM, 2010.

[17] Albert Gatt and Ehud Reiter. Simplenlg: A realisation engine for practical applications. In *Proceedings of the 12th European Workshop on Natural Language Generation (ENLG 2009)*, pages 90–93, 2009.

[18] Fatih Gedikli, Dietmar Jannach, and Mouzhi Ge. How should i explain? a comparison of different explanation types for recommender systems. *International Journal of Human-Computer Studies*, 72(4):367–382, 2014.

[19] NA Herk and RC Thompson. International technical brief for the thomas-kilmann conflict mode instrument. *CPP Inc*, 2011.

[20] Jonathan L Herlocker, Joseph A Konstan, and John Riedl. Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 241–250. ACM, 2000.

[21] Daniel Herzog. Recommending a sequence of points of interest to a group of users in a mobile context. In *Proceedings of the eleventh ACM conference on recommender systems*, pages 402–406. ACM, 2017.

[22] Daniel Herzog, Christopher Laß, and Wolfgang Wörndl. Tourrec: a tourist trip recommender system for individuals and groups. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 496–497. ACM, 2018.

[23] Daniel Herzog and Wolfgang Wörndl. A user study on groups interacting with tourist trip recommender systems in public spaces. In *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, pages 130–138. ACM, 2019.

[24] Anthony Jameson. More than the sum of its members: challenges for group recommender systems. In *Proceedings of the working conference on Advanced visual interfaces*, pages 48–54. ACM, 2004.

[25] Anthony Jameson, Stephan Baldes, and Thomas Kleinbauer. Two methods for enhancing mutual awareness in a group recommender system. In *Proceedings of the working conference on Advanced visual interfaces*, pages 447–449. ACM, 2004.

[26] Anthony Jameson and Barry Smyth. Recommendation to groups. In *The adaptive web*, pages 596–627. Springer, 2007.

[27] Öykü Kapcak, Simone Spagnoli, Vincent Robbemond, Soumitri Vadali, Shabnam Najafian, Nava Tintarev, Julia Neidhardt, Wolfgang Wörndl, Tsvi Kuflik, and Markus Zanker. Tourexplain: A crowdsourcing pipeline for generating explanations for groups of tourists. 2018.

[28] Ralph H Kilmann and Kenneth W Thomas. Developing a forced-choice measure of conflict-handling behavior: The" mode" instrument. *Educational and psychological measurement*, 37(2):309–325, 1977.

[29] Judith Masthoff. Group modeling: Selecting a sequence of television items to suit a group of viewers. In *Personalized digital television*, pages 93–141. Springer, 2004.

[30] Judith Masthoff. Group recommender systems: Combining individual models. In *Recommender systems handbook*, pages 677–702. Springer, 2011.

[31] Judith Masthoff. Group recommender systems: aggregation, satisfaction and group attributes. In *recommender systems handbook*, pages 743–776. Springer, 2015.

[32] Judith Masthoff and Albert Gatt. In pursuit of satisfaction and the prevention of embarrassment: affective state in group recommender systems. *User Modeling and User-Adapted Interaction*, 16(3-4):281–319, 2006.

[33] Kevin McCarthy, James Reilly, Lorraine McGinty, and Barry Smyth. Thinking positively-explanatory feedback for conversational recommender systems. In *Proceedings of the European Conference on Case-Based Reasoning (ECCBR-04) Explanation Workshop*, pages 115–124, 2004.

[34] Will S Monroe. Social organization: A study of the larger mind. *The Journal of Philosophy, Psychology and Scientific Methods*, 7(2):50–51, 1910.

[35] Shabnam Najafian and Nava Tintarev. Generating consensus explanations for group recommendations: an exploratory study. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*, pages 245–250. ACM, 2018.

[36] T Nguyen and Francesco Ricci. Supporting group decision making with recommendations and explanations. CEUR-WS, 2016.

[37] Lara Quijano-Sanchez, Juan A Recio-Garcia, Belen Diaz-Agudo, and Guillermo Jimenez-Diaz. Social factors in group recommender systems. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 4(1):8, 2013.

[38] Lara Quijano-Sanchez, Christian Sauer, Juan A Recio-Garcia, and Belen Diaz-Agudo. Make it personal: a social explanation system applied to group recommendations. *Expert Systems with Applications*, 76:36–48, 2017.

[39] Ehud Reiter and Robert Dale. Building applied natural language generation systems. *Natural Language Engineering*, 3(1):57–87, 1997.

[40] Shapira B Ricci F., Rokach L. (2015) recommender systems: Introduction and challenges. In *Ricci F., Rokach L., Shapira B. (eds) Recommender Systems Handbook. Springer Boston, MA*.

[41] Christophe Senot, Dimitre Kostadinov, Makram Bouzid, Jérôme Picault, Armen Aghasaryan, and Cédric Bernier. Analysis of strategies for building group profiles. In *International Conference on User Modeling, Adaptation, and Personalization*, pages 40–51. Springer, 2010.

[42] Rashmi Sinha and Kirsten Swearingen. The role of transparency in recommender systems. In *CHI'02 extended abstracts on Human factors in computing systems*, pages 830–831. ACM, 2002.

[43] Kenneth W Thomas. Thomas-kilmann conflict mode. *TKI Profile and Interpretive Report*, pages 1–11, 2008.

[44] Nava Tintarev. Explaining recommendations. In *International Conference on User Modeling*, pages 470–474. Springer, 2007.

[45] Nava Tintarev and Judith Masthoff. A survey of explanations in recommender systems. In *Data Engineering Workshop, 2007 IEEE 23rd International Conference on*, pages 801–810. IEEE, 2007.

[46] Marko Tkalcic, Amra Delic, and Alexander Felfernigab. Personality, emotions, and group dynamics, 2018.

[47] Thi Ngoc Trang Tran, Müslüm Atas, Alexander Felfernig, Viet Man Le, Ralph Samer, and Martin Stettinger. Towards social choice-based explanations in group recommender systems. In *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, pages 13–21. ACM, 2019.

[48] Jesse Vig, Shilad Sen, and John Riedl. Tagsplanations: explaining recommendations using tags. In *Proceedings of the 14th international conference on Intelligent user interfaces*, pages 47–56. ACM, 2009.

Appendix A

Generating Explanations

We have generated the explanations designed by us using Natural language generation (NLG) and SimpleNLG. In this section, a brief discussion, on how NLG was used to generate explanations is given. The synthetic numerical ratings given for the candidate items (places) (refer Figure 3.1) are converted into text using logic and template-based NLG (Simple NLG) [17].

A.1 Tasks in NLG

"Natural Language Generation is a sub-field of computational linguistics and artificial intelligence" [39]. In addition to designing and evaluating explanations, I have generated them using NLG.

The first step is to analyze the **requirements** of the text that has to be generated. For this, the initial corpus of texts are gathered from the various sources available and a "target" corpus of text to be generated is finalized. The literature study gave us a number of ways in which social explanations and explanations without social components are designed. *Therefore the literature is the initial corpus from which we finalized our target corpus after preliminary study 1 and study 2.* This content determined is then segregated as directly available text, unchanging text, computable data, and unavailable data [39]. This is explained in the next paragraph with an example from explanations.

For example, consider a part of the transparent explanation as mentioned in chapter 3 *".. We note that place C is rated highly by you, but it has not been included in the sequence. This is because place C is rated low by Mary and John... Although you are recommended the places you don't like, you get to visit places you love first. Furthermore, almost all group members, are compromising in similar situations of disagreement like these to satisfy the group".* The assignment of parts of the explanation is listed below.

- Unchanging text is " we note that", "but it has not been included in the sequence", "this is because ".
- Directly available text would be "place C is rated highly by you".
- Computable data is "place C is rated low by Mary and John".

- Unavailable data would be "Although you are recommended the places you don't like, you get to visit places you love first. Furthermore, almost all group members, are compromising in similar situations of disagreement like these to satisfy the group".

After requirement analysis, the next step is to do **text planning and sentence planning** before the final task of linguistic realization. There are six tasks that could be done in order to create text. They are Content Determination, Discourse Planning, Sentence Aggregation, Lexicalization, Referring Expression generation and, Linguistic Realization [39].

We have considered content determination and discourse planning for this thesis. But, we have not performed Sentence Aggregation, Lexicalization and Referring Expression generation in planning sentences. This is because we are going to create templates and generate text from synthetic ratings which will be placed in-between templates. For example, consider the explanation, "We note that *place C is rated highly by you*, but it has not been included in the sequence. This is because *place C is rated low by Mary and John*". The italicized phrases are generated from synthetic ratings using simple NLG[17] while other sentences act as **templates**. Finally the sentences are linguistically realised [39].

In content determination, the text message is represented as "entities, concepts, and relations" [39] . In this thesis,

- Entities are place names, rating as low-mid-high, names of people;
- Concepts are a sequence of place names whereas,
- Relations are the relation between rating and place names, people.

This concept is applied to each of the sentences of the explanation that resulted at the end of preliminary study 2. In Discourse planning, the structure of the sentence is planned with begin-middle-end. For this, each planned phrase of the explanation is linked with other prefix and suffix phrases to create a tree structure. This tree structure results in a final text that we intend to generate. Content determination and Discourse planning help for text planning.

A.2 SimpleNLG

SimpleNLG is a "realisation engine" (library) built by Albert Gatt and Ehud Reiter [17]. The code is provided in JAVA object-oriented programming language. The jar files of SimpleNLG is used for syntactic, morphological and orthographic correct sentences. We have used the current release of SimpleNLG is V4.4.8 [2]. The "lexicon" object is created from the lexicon java class. This class has all the variables and features mentioned in Figure A.1. Then an object of the java class "NLGFactory" and object of the java class "Realiser" is created. Both the latter objects use the database available in lexicon class programmatically [17]. From the NLGFactory class, we have used SPhraseSpec, CoordinatePhrase, NPPhraseSpec classes to access their respective functions in order to generate explanations. A detailed overview of how to use these classes and functions is given in [2].

	Feature	Values	Applicable classes
lexical	ADJPOSITION	Attrib$_{1/2/3}$, PostNominal, Predicative	ADJ
	ADVPOSITION	Sentential, PostVerbal, Verbal	ADV
	AGRTYPE	Count, Mass, Group, Inv-Pl, Inv-Sg	N
	COMPLTYPE	AdjP, AdvP, B-Inf, WhFin, WhInf, ...	V
	VTYPE	Aux, Main, Modal	V
phrasal	FUNCTION	Subject, Obj, I-Obj, Prep-Obj, Modifier	all
	SFORM	B-Inf, Gerund, Imper, Inf, Subj	S
	INTERROGTYPE	Yes/No, How, What, ...	S
	NUMBERAGR	Plural, Singular	NP
	TENSE	Pres, Past, Fut	VP
	TAXIS (boolean)	`true` (=perfective), `false`	VP
	POSSESSIVE (boolean)	`true` (=possessive), `false`	NP
	PASSIVE (boolean)	`true`, `false`	VP

Figure A.1: Variables and features available in SimpleNLG [17]

A.3 Logic for ratings to text

The explanation used in the final evaluation had a structure that was the sandwich model[11] of the explanation. The explanation will provide an information in the following order.

1. Algorithmic Transparency
2. Positive Reassuring Explanation about "Places you like is recommended"
3. Repair Related Explanation about "Places you like is never recommended at all"
4. Repair Related Explanation about "Places you like is not recommended in the preferred order"
5. Repair Related Explanation about "Places you dislike is recommended"
6. Positive Assuring explanation about "People who also had to compromise".

This order is applied to both transparent and privacy preserving explanations. We have created a java function for each of the six points above which uses template based NLG and logic. A very detailed explanation of how I have implemented the logic and Java code is available in Github [1].