ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
Human Computer Interaction Group

PROJECT REPORT

EYE-TRACKING GROUP INFLUENCE
- EXPERIMENT DESIGN AND RESULTS -

Coordinator teacher:
Dr. Pearl Pu

PhD Student:
George Popescu

Switzerland, 26/09/2012

Table of Contents

Abstract

The decision success of a group is strongly related to its members' satisfaction while adapting to common outcomes which, in turn, relate to familiarity and trust. Evaluation of determinant factors enhancing group dynamics is essential for understanding social influence in a group setting.

This report presents the design and results of a group influence study in a music scenario, in which people share their tastes and participate in group decisions. We designed and ran an experiment with the purpose of tracking group influence in our GroupFun music recommender system. Specifically, we were interested in measuring the extent to which people change and align their opinions when facing various individual preferences coming from their group members. We describe the design, implementation and results of our between-groups user study. Two groups of 9 and 10 members are evaluated one against the other by means of comparing decision acceptance, participant satisfaction and decision changes using questionnaire and eye-tracking data.

In our study we first asked participants to give individual ratings to songs uploaded to their group and then to perform an eye-tracking task in which they saw their peers' names and ratings and were asked to make a rating decision again. The eye-tracking outputs we used present clear associations between various areas of the interests while fixation times are useful to understand group influence relative to closest members.

Our results focus on group conformity, or the degree to which a person changed his own rating aligning to the group decision. We prove that the stronger social relationships through familiarity and trust are in a group, the more individuals adapt to the group decision and are more inclined to change their preferences. The characteristics of the group, namely member familiarity and trust, dictate the overall group satisfaction and outcome decision acceptance. In the more connected group participants not only change their preferences more being more influenced by other members, but they are also more satisfied with the overall outcome. Also, the gaze correspondences users produce between other members' names and their ratings represent a good indicator for understanding how individuals evaluate a given song.

In summary, we present an in-depth, between-group analysis incorporating preference dynamics which we correlate with eye-tracking data to understand decision making induced by social influence. The results mentioned in this report have strong implications for decision making strategies and group recommendation techniques as well as for psychology and personality research. More experiments should continue this emerging and promising area of research.

Author Keywords: Area of Interest, Change, Eye-gaze, Eye-tracking, Familiarity, Group, Interface, Music, Ratings, Recommender System, Social Influence, Songs, Trust.

ACM Classification Keywords: K.4.0 Computers and Society: General. H.5.2 [Information Interfaces and Presentation (e.g. HCI)]: User Interfaces – theory and methods, evaluation/methodology, user-centered design.

General Terms: Design, Human Factors, Measurement

1. Introduction

Imagine you and your friends have to decide the location for your upcoming holidays, new activities that you would like to plan together, or the playlist songs for a party you are all going to attend. How would you participate? Would your preferences be strict or will you adapt to the group decision? Which members in your group do you know best? Which other members do you trust most? For which of them do you know or can you guess their music preferences? Would you like to discover such preferences? A group recommender system helps your group make a qualitative decision together!

Whether planning the ideal holiday trip from an overwhelming list of destinations, selecting a best movie to watch with your friends or choosing entertaining music for carpooling every day to work, individual satisfaction is essential for accepting group decisions. Preference aggregation systems contain implementation of algorithms which aim at maximizing the overall group satisfaction [1, 4, 5, 21, 25, 26] or providing an intermediate common output that the group would agree or improve upon [2, 11, 20, 23]. The state-of-the-art scientific literature in the field has pursued two research directions: one which deals with group dynamics [20, 22, 25, 29 and 33] and the other which analyses decision accuracy and user acceptance [4, 10, 23, 31 and 32].

Researchers use tools to describe and understand decision behavior in both online and off-line environments for specific domains such as apartment search, music, games, etc. Nevertheless, group dynamics was correlated with individual and group performance [22, 23 and 29]. With the emergence of social networks various visualization techniques are able to capturing the temporal group interactions [22].

Unfortunately, little is known about how the group influences the individual (and vice-versa) in online systems and how it contributes to his satisfaction and decision change. Consequently, large-scale network studies have the advantage of presenting an overall understanding of social network dynamics but are limited by an in-depth comprehension of factors which determine decision change at an individual level [35]. Nevertheless, trust-based group models provide a reliable framework for understanding individual decision making [2, 19 and 37].

Traditionally, group recommender systems (GRS) have evolved by developing individual preference specification models which correspond to users interests [4, 5, 11, 20, and 31]. However, these preferences change significantly over time and are context dependent: e.g. a group of friends would like to listen to one type of music for Saturday night’s party and of different type while carpooling to work together on Monday morning. In addition, preferences may vary according to the group: one can share some music preferences while when with closest friends and different ones while when with family or strangers. Thus, GRSs propose new algorithmic and design approaches by: incorporating social relationship interactions [20] and combining user and item information together [4, 5]. Personalized recommendations have the key advantage of removing popular, uninteresting, recommendations and favoring the discovery of new items [10, 25] even within the same group.

One significant limitation in this field is that the development of algorithms with increased accuracy in knowing users’ tastes does not necessarily constitute a silver bullet or key factor for improving recommendations. Generally, such algorithms demand for an extensive user effort and produce only a small increase in user acceptance [31, 25]. Other studies focus on the “pursuit of satisfaction” in GRS instead and prove the role of an affective state in decision making and individual satisfaction [23].

Eye-tracking devices represent a precise and reliable technology measuring users’ eye movements facing an interface. Such technology has been used for understanding individual decision making, for example when people look at various items with the aim of identifying objectively and indirectly the features that are of interest for an individual or customer segment. Such technology was not previously used in the study of group influence but only in online marketing related studies for which mainly retailers assessed website features capturing the visitors’ interest through their eye-gaze for a specific item and/or recommendation technique [7, 8, 9 and 12]. Finally, social influence analysis was addressed in large-scale online networks with the purpose of identifying the most influential person(s) and understanding the influent-follower role [16, 19 and 35]. In was observed that people switch their preference when receiving recommendation from friends or people they know well conforming to the group choice [36, 37 and 38].

The human decision mechanism represents a complex system and is influenced by numerous factors which we aim to identify through a group-based user study. More precisely, we investigate the issue of group influence in a music context for a group to select desired songs for a party they plan and attend together. This issue is generally known as the preference aggregation problem [1, 2 and 21] in which multiple individual tastes get combined by a system which recommends best options to the whole group.

We designed a group influence experiment in which we aimed at tracking and quantifying people's decision making through the use of two systems: the GroupFun group music recommender system, in which participants submit individual preferences, and an eye-tracking-based set of interfaces in which subjects are exposed to other group members' preferences. As mentioned before, the group decision making area of research lacks experiments and analysis of social influence sustained by both rating/preference submission and eye-tracking information. Therefore, correlating subjective a priori known evaluation group factors, such as familiarity and trust, with participants' objective eye-gaze, yield interesting results for both the social science and group decision research fields. In the present study we draw correlations between the two groups' connected networks' properties and define similarity and group influence metrics.

Our results show that social influence is stronger in groups in which participants are most familiar with and trust their peers' tastes. Moreover, we found trust to be a better measure than familiarity for predicting people eye-gaze behavior. Most importantly, group conformity is correlated with group connectivity and peer pressure. Thus, the experiment sustains social conformity theory which states that people align their preferences to those of the group.

The specific contributions of this paper are:

(1) The design of a between-groups experiment combining subjective information consisting in familiarity and trust values with objective observations given by eye-gaze correspondences and focus times. We briefly include the main features of our group music recommender system that participants used to contribute music to their group and submit individual ratings. For the experiment we present the most

important statistical results across all conditions, such as: consensus vs. divergence, trust vs. popularity and group size (2 closest, 5 closest and all other members).

(2) A comparison between the two groups' characteristics. We show that the more connected the group is, i.e. the denser the familiarity and trust networks are, the higher the change rate is. In other words people adapt more to the group decision by changing their own evaluation when social bounds are stronger. We present an extensive group satisfaction comparison as well as rating change statistics together with preference correlation values.

(3) The representation of an overlay between the subjective values from each group computed as normalized familiarity and trust and the objective eye-gaze correspondences per participant aimed at highlighting the most important members that influence the current participant's decision. We map the two values one to another and report that the eye-tracking data represents a reliable source of information for constructing a group's social network with high accuracy. By opposition, the music correlation values between participants do not represent a reliable source for deriving social bonds. In other words, in our study, the few individuals who knew and trusted others most did not actually have the same music taste. As a general consequence, the objective data gathered from the eye-tracker strengthen the social alignment results based on familiarity and trust subjective information.

Additionally, we investigated if individuals follow the "average" group ratings, some specific members of the group or only closest members. Our data shows that individuals do not follow pre-defined rating schemas but rely on a mixture of information presented to them incorporating their previous preferences and others' ratings.

The remainder of this report is structured as follows. First, we summarize related work in the field from two research areas: social science and influence and eye-tracking techniques. Then, we describe the experiment setup in detail presenting all structural elements. We continue by emphasizing the 3 building blocks of our study: an online questionnaire, the GroupFun system and the eye—tracking experiment. We next discuss extensively the main results and their implications focusing on certain advantages and limitations. Finally, we summarize our ideas through the conclusions and future work.

2. Related Work

There are two research directions which are mainly related to the current report. The first one is represented by psychological studies related to social influence [6, 15, 24, 27, 34, 35 and 36] and persuasiveness [28, 32] and the second one represented by the role of eye-tracking technology to as a means of objective evaluation of information [7, 8, 9 and 30].

2.1 Social Influence

Group Recommender Systems scientific literature has recently focused on the development of persuasion methods and techniques that can contribute to one's satisfaction or well-being when supported by others. Fit4Life [32] takes the challenge of incorporating design guidelines from recent research under one system design with the purpose of promoting healthy behavior and ideal weight. Despite the fact that the authors include only general lines of system description, numerous advantages can be foreseen with respect to the use of such persuasive technology.

Under the "social pressure" of others exercising and posting activity reports on their social network's wall, individuals can benefit from adopting a healthier life-style [32]. The authors introduce rewards and motivations based on persuasive methodologies to encourage behaviors that address the social problem presented by obesity. The persuaders are positioned at two distinct levels: at the first one the system designers send messages to users to motivate them to use the system and, on the other, the others subscribers act through social networks influencing each user "to do what is best for them".

Similar to the approach from Fit4Life, in our eye-tracking experiment [27] we set the persuaders to be the other group members through their names and our experiment design through controlled study conditions and ratings. Another key similarity between our group recommender system, GroupFun [26], and Fit4Life [32] is that both are implemented as Facebook connectors. However, instead of broadcasting rating and listening habits to one's wall with respect to users' practice and habits as Fit4Life does, our system was designed to only gather rating information favoring internal decision context and group privacy.

Psychology literature additionally describes experiments and methods that enhance users' participation given others' influence. In 2 experiments conceived by Sukuran et al. [34] individual thoughtful effort is tested with respect to others' similar behavior. Specifically, taking the context of online participant commenting, the authors analyze the length of submitted comments, the time taken to write them and the number of issue-relevant thoughts that they contain and prove that subjects conform to high vs. low norms of thoughtfulness. Subjective perceptions which people form about online situations they take part in are modeled through interfaces and web-site design.

Similar studies found that online contributions can be motivated by the individual's perception of his level of participation relative to others [14] or degree of identification with the social group [33]. Behavioral influences can be traced to social norms which are inferred from what others are doing [16, 24, 33 and 34].

Social influence studies represent a main area of research in psychology literature [15, 16 and 36]. With the emergence of the internet and group decision technology some experiments focus on social influence in a computer-mediated environment in which participants rely on technology to produce a group decision [19, 28 and 37].

Postmes et al. [28] investigate the role of anonymity in group behavior and find out that by protecting one's identity certain benefits favor group decision through transparent decision rules. Some other studies tackle the problem of human decision in a group setting [25, 26, 29 and 33] pointing to factors which are fundamental for social influence while others emphasize that social influence accounts for a large proportion of human behavior and that decision making is under-evaluated [24].

2.2 Eye-Tracking

The other side of academic work which is strongly correlated with the research results included in the current report covers eye-tracking technology employed to capture highly important data revealing one's decision making process.

Traditionally, eye-tracking studies have focused on usability evaluations of products or websites aiming at capturing users' interest for certain features [31]. Numerous guidelines have been proposed for designers of software systems need to adapt their work to the way people look at certain features on their website, such as organizational interfaces [8, 9].

There are two general interests in the eye-tracking community. On the first hand, eye-tracking studies focus on the analysis of the gaze it-self with the purpose of determining efficient clustering methods for both gaze plots and fixation times [17]. The methods included in such studies are particularly interesting for the analysis of large eye-tracking output gaze-plots and raw data. On the other hand, scientists are concerned about human decision making and thinking trying to grasp the main visual factors which govern judgment.

In an eye-tracking study Pretorius et al. [30] compare how adults and children learn to use unfamiliar games. The experiment reveals important insights on both children's and adult's gaze patterns. Another related study by Birkett et al. [3] proves that eye-movements represent a reliable measure for understanding web engagement in the case of young children. Also, another recent study by Dumais et al. [18] reveals differences in individual eye-gaze web browsing and search.

More importantly, in the context of recommender systems, eye tracking technology proved to be a reliable technology for identifying users' interest in using the system as well as analyzing the efficiency of a recommendation process [8]. This technology can be employed for the use recommender systems helping a group to decide on the music they would listen to [26].

Other related eye-tracking studies focus on following user behavior in recommender interfaces with the goal of understanding certain psychological aspects that model one's decision process [13]. Similar experiments have identified certain consumer decision patterns through eye-gaze analysis [9] which are fundamental for the design of better recommendation systems. Under social influence individuals act differently in low-preference compared with high-preference domains willing to adapt to others decisions. Thus, eye-tracking technology can be used to analyze decision patterns in groups, yielding results which help explain the change in one's own judgment.

3. Experiment Design

The experiment design was based on the following preference elicitation scenario: rating and listening to music alone in front of a computer and re-submitting ratings after seeing clear interfaces with others' decision. The main goal was to collect the highest amount of useful subjective and objective information relevant for social influence analysis. Both the GroupFun system and the eye-tracking interfaces were conceptualized to be easy to use and facilitate participants' understanding of the information displayed to them.

3.1 Participants

19 participants split into 2 groups where recruited to participate in the group influence experiment. Their ages range from 23 to 30, and were mainly university students or employees. Subjects had 12 different national backgrounds: Switzerland, France, China, India, Russia, Romania, Italy, etc. and 3 different educational backgrounds: Master, PhD and Post-Doc students.

The first 9 participants formed Group1 and the last 10 formed Group2. In the first group there were only male participants whereas in the second there were 6 male and 4 female members. One important aspect in group structure is that members in each group knew all other members but not the members from the other group. More precisely, Group1 members were students in the "Conducting User Studies" course offered at EPFL during fall 2011 whereas Group2 members were Master, PhD and Postdocs working closely on their research in two close labs: Human Computer Interaction and Artificial Intelligence.

Subjects only interacted with the other peers in their own group, but not the other. Thus the results obtained for each group are not influenced by the results obtained for the other.

3.2 Duration

The study was carried out over a period of 3 weeks. In the first week the two groups used the GroupFun music recommender system to create a group, upload music and give ratings individually, without any group interface and influence. During weeks 2 and 3, eye-tracking experiments were carried out with the purpose of recording users' eye-gaze and analyze their newly updated ratings by comparing them with previous ones. In week 2, all members of Group1 performed the eye-tracking phase of the experiment whereas in week 3 all members of Group2 followed the same tasks. In total we collected more than 5 hours of eye-tracking gaze data and had rating information for 86 songs.

3.3 Reward

The study proposed a music incentive. The reward that all users received at the end of the experiment was a collection of all group songs uploaded by all group members (appendix). We created an archive for each group and sent the music library to all members of the same group. We sent 39 contributed songs to all Group1 members and 47 contributed songs to all Group2 members. Additionally we shared similarity music profiles and rating correlation values with some members who expressed their desire to know such information.

3.4 User Tasks

Our experiment's main user task was to "choose songs for a party that all group members would organize together".

We aim to compare group satisfaction based on reported ratings under individual preference specification against that under social influence-based interfaces. We further analyze the role of social influence in decision dynamics on the grounds of relative rating change.

The detailed steps that each participant chronologically pursued are depicted in Figure 1.

Subjects were asked to perform the 3 following tasks:

(1) Complete an online questionnaire evaluating familiarity and trust with respect to other group members. For assessing familiarity and trust subjects used ratings from "1" to "5" for each other group members. Both familiarity and trust networks per group were generated based on this data (Figure 3).

(2) Use GroupFun[1] to create a group, upload music and give individual ratings to group songs. Users logged in with their Facebook accounts to GroupFun and joined the virtual group corresponding to the real one. Then they uploaded several .mp3 files which the system displayed in an individual preference elicitation interface. Using this interface users were able to listen to the songs, fast forward, pause and give a rating from "1" to "5" to any number of contributed songs in the group.

(3) Listen and give ratings to 24 group songs while seated in front of group interfaces containing group members' names and their ratings and being recorded by an eye-tracking device. All 24 songs were extracted from the total group contribution, i.e. each song belonged to one of the group members.

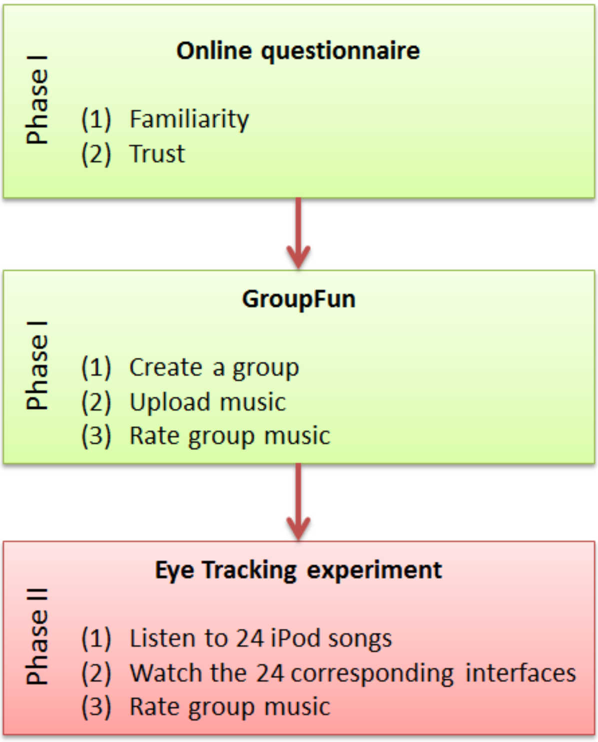


Figure 1. Experiment setup and phases

Tasks 1 and 2 represent Phase I (Figure 1, in green) and were done by each group during the first week. Task 3 represent Phase 2 (Figure 1, in red) and was done by each group subsequently during weeks 2 and 3. In Phase I subjects first submitted familiarity and

[1] http://grpupc1.epfl.ch/~george/GroupFun/

trust information. Then they used GroupFun and created Group1 named “Cool party” and Group2 named “Legendary party”, uploaded several favorite songs to the group and gave ratings individually, without any external influence from other members. In Phase 2 the role of the eye-tracker was to gain in-depth objective data about group decision.

The two main conditions being investigated are: 1) individual preference elicitation which represents initial ratings given my each member and 2) group influence preference elicitation which consists in ratings given during the eye-tracking task.

3.5 Objectives

In a social context people adapt their decision to that of the group. We expect our volunteers to change decision given their group orientation rather that keeping strict previously-stated preferences through ratings. Our main objective is to analyze group influence, define metrics for analysis and compare different types of users in order to understand which factors produce social alignment and which others push users towards the reconfirmation of their preferences. Furthermore, we draw correlations between the implicit familiarity and trust network information and explicit eye-tracking data.

Given the specific characteristics of each group we aim at analyzing how the familiarity and trust networks influence each group’s preference dynamics. More specifically we analyze group satisfaction before and after the eye tracking session and draw correlations among music tastes. We expect that the higher group connectivity yields greater social influence and rating changes among group members. Conversely, the more distant members feel inside a group the more self-oriented they are expected to be and not change their ratings reaching the same satisfaction under both the individual setting and the group influence conditions.

We also compute rating correlations between members of the same group in order to determine closest members preference-wise and compare this objective metric with the subjective evaluation given by our users after answering the online questionnaire. We also correlate social bounds between group members with data obtained from the eye-tracking device to understand specific group factors that model individual decision. Overall, our main goals are: (1) to understand how the group characteristics model individual decision making and (2) to compare members of two different groups among them-selves in terms of familiarity, trust, preference, group satisfaction, eye-gaze and rating change.

3.6 Materials

During Phase I (Figure 1) volunteers used an online form to fill in their responses. The link to the form was provided by e-mail so that respondents could fill in their answers privately at any time. Then, they proceeded to use the GroupFun music recommender system[2] for creating a group, uploading music and contributing to the group decision via a self-evaluation of group songs through ratings.

In addition to this, we used the Tobii 1750 eye-tracking device[3] consisting in a monitor with embedded cameras based on infrared light to capture pupil movement. We attached a desktop computer running Windows XP and the predefined ClearView 2.7.1 software capable of capturing users' eye gaze, display fixation points and generate heat-maps. After an initial calibration for each participant, we asked our users to look at the screen in a natural and comfortable way with the help of a head mounted object.

The second device for our experiment was an Apple iPod Shuffle[4] on which we recorded 24 recommended songs played for each group participant in the interface order displayed on the eye tracker. Members of each group had the same 24-songs playlist recorded in the iPod but songs per group different according to each group's specific contribution.

Screenshots with our materials are included in the appendix section of the report.

[2] http://grpupc1.epfl.ch/~george/GroupFun/

[3] http://grpupc1.epfl.ch/~george/Tobii1750.gif

[4] http://grpupc1.epfl.ch/~george/iPod%20shuffle.jpg

4. Online Questionnaire

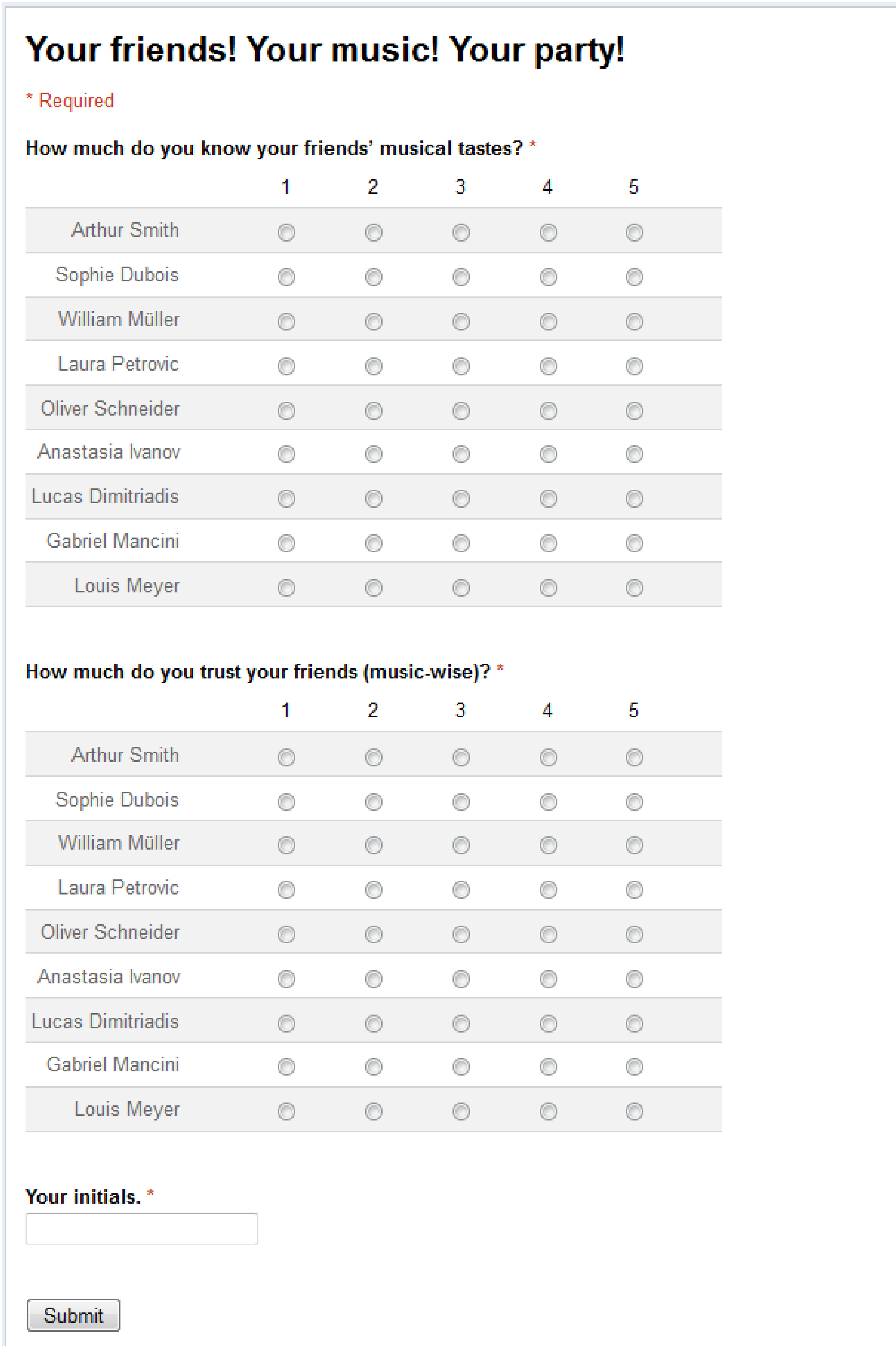
Before the experiment we asked our participants to fill out an assessment questionnaire[5]. In both Figure 2 and at the link provided at the footnote of this page we included a copy of the real questionnaire replacing our participants' real names with others preserving their privacy and name information.

Two questions were used to assess familiarity and trust information: (1) "How much do you know your friends' musical tastes?" and (2): "How much do you trust your friends (music-wise)?" Both questions were mandatory and each participant had to submit a rating value for each of his peers. The first one assesses self-revealed familiarity whereas the second one self-revealed trust on a 5 point Likert scale. We asked our participants to submit any value from "1" to "5" for their peers. The completion time of the questionnaire lasted around 2 minutes per participant.

The role of the two above questions is double-folded. First, we explore the groups' internal subjective network (familiarity and trust) structure as seen through the eyes of its members. Secondly, we map this subjective information with (later-collected) eye-tracking data to see to what extent the familiarity and trust values correspond to each participant's actual behavior. This represents a distinctive trait of our experiment compared with similar approaches in the scientific literature. We engage in comparing subjective and objective parameters which model group behavior yielding results which are useful for the understanding of group decision making and social influence.

In our music setting, we hypothesize that people are more inclined to trust others' even though they do not know much of others' music tastes in advance. We expect that the music discovery process would allow them to enhance group trust over familiarity. We also acknowledge the fact that in other contexts, familiarity may dominate trust instead.

5 https://docs.google.com/spreadsheet/viewform?formkey=dGRjbjQyRVZzYnk0XzFDeHk2aVV6dWc6MA#gid=0



Your friends! Your music! Your party!

* Required

How much do you know your friends' musical tastes? *

	1	2	3	4	5
Arthur Smith					
Sophie Dubois					
William Müller					
Laura Petrovic					
Oliver Schneider					
Anastasia Ivanov					
Lucas Dimitriadis					
Gabriel Mancini					
Louis Meyer					

How much do you trust your friends (music-wise)? *

	1	2	3	4	5
Arthur Smith					
Sophie Dubois					
William Müller					
Laura Petrovic					
Oliver Schneider					
Anastasia Ivanov					
Lucas Dimitriadis					
Gabriel Mancini					
Louis Meyer					

Your initials. *

Submit

Figure 2. Familiarity and trust assessment online questionnaire

4.1 Familiarity

When analyzing the first variable, familiarity, we noticed two categories of users in both groups: one includes members who are more isolated and the other includes people who connect with others. In the latter case, familiarity distribution covers most group members. A visualization graph conceptualized in Gephi[6] (Figure 3) uses edge weights and positioning to identify the persons who are most influential in each group. We designed an oriented graph (Figure 3) in which the greater the weight given by Pi to Pj the thicker the directed line connecting them.

In Table I below we include the familiarity values submitted by each participant with respect to the other participants in the same group.

There are few observations which can be immediately inferred from the two tables. The first one is that, in replying to the first question stated in the online questionnaire: “How much do you know your friends’ musical tastes?” Most users gave few rating values different from the minimum, which is 1. The second observation is that most users did not consider a normalization process when ratings, e.g. P17 in the second group gave ratings of 3 to all his peers, P8 in the first group gave only ratings of 1 to all of his peers. According to these ratings, the difference between the two participants is that P8 estimated that he does not know anything about others whereas P17 stated that he knows to some extent all others’ music tastes. Thus, it is difficult to quantify the rating differences among participants. Finally, the third observation is that the rating behavior of our participants is quite different: if we compare P2 with P7, for instance, we notice that P2 used a wider palette of ratings to evaluate how much he knows others’ music tastes whereas P7 used only 2 ratings of “2” to other 2 group members without using any higher values.

Overall we consider the familiarity values (and the graphs based on these values) as an approximation of the groups’ social network with respect to member’s interactions. Despite the general data sparsity which can be observed for all participants we remark that the second group is more connected than the first one, familiarity-wise.

---

[6] http://www.gephi.org/

Average familiarity in Group1 is 1.80 (out of 5) and average standard deviation is 0.81. The situation is slightly different in Group2 where both average familiarity and average standard deviation are lower: 1.59 and 0.55 respectively. Furthermore, there are one participant in Group1 (P8) and 3 participants in Group2 (P15, P16 and P17) who submitted the same familiarity ratings for all other members, thus yielding a standard deviation equal to 0. P8, P15 and P16 submitted all ratings of “1” and P17 all ratings of “3”. We notice that P2 and P3 are most familiar with others’ music tastes (average familiarity of 3.8 and 2.6, respectively) in Group1 and P17 and P14 (average familiarity of 3.0 and 2.3, respectively) in Group2. To conclude this section we report that both average and standard deviation values are small for the first factor, familiarity.

Table I. Familiarity values for Group1 (top) and Group2 (bottom)

Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9	Average	Std.
P1		1	1	2	1	1	1	1	1	**1.1**	0.35
P2	1		5	3	4	5	5	3	4	**3.8**	0.39
P3	3	5		2	3	1	2	3	2	**2.6**	1.19
P4	3	1	1		1	1	1	1	1	**1.3**	0.71
P5	1	1	1	1		2	4	3	1	**1.8**	1.16
P6	1	1	1	1	3		3	2	3	**1.9**	0.99
P7	1	1	1	1	2	1		1	2	**1.3**	0.46
P8	1	1	1	1	1	1	1		1	**1.0**	0.00
P9	1	1	1	1	1	1	4	1		**1.4**	1.06

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Average	Std.
P10		1	1	1	1	1	1	4	1	1	**1.3**	1.00
P11	3		1	1	1	2	1	3	2	3	**1.9**	0.93
P12	1	1		2	1	3	1	1	2	1	**1.4**	0.73
P13	1	1	3		1	1	1	1	1	1	**1.2**	0.67
P14	4	2	2	2		2	3	1	2	3	**2.3**	0.87
P15	1	1	1	1	1		1	1	1	1	**1.0**	0.00
P16	1	1	1	1	1	1		1	1	1	**1.0**	0.00
P17	3	3	3	3	3	3	3		3	3	**3.0**	0.00
P18	1	1	1	1	1	1	1	1		2	**1.1**	0.33
P19	2	2	1	1	1	1	1	2	4		**1.7**	1.00

4.2 Trust

The second variable we analyzed is trust. The responses submitted by the first group's members demonstrate that people who know others' music tastes trust them more for music recommendation: P1 and P4 trust themselves more than others and trust others' tastes even though they don't know them. However, members who do not know their peers' tastes are tempted to trust them given their tight social relationships and openness to discover new music: P2 and P3 trust other members' preferences. The trust distribution covers most group members. Similar results yield for the second group.

The average of trust values per participant in both groups show higher values that for familiarity. It's interesting to compare the 2 familiarity tables from the previous section with the 2 trust tables from below. First, we notice participants that are isolated in their groups. It is the case of participant P15 who submitted only "1"-values for both familiarity and trust. Others, such as P16 and P17 submitted only "3"-values for trust. P17 submitted exactly the same values for familiarity whereas P16 upgraded his ratings from "1"s to "3"s from familiarity to trust.

Another interesting phenomenon is that on average the standard deviation of ratings varies only slightly from familiarity to trust in each group. Moreover, the second group is the more connected one as higher ratings are mapped from the 10 participants in this group whereas generally, in the first group, the 9 participants know and trust only a few other members on average.

In Group1 average trust for all 9 members is 2.21 (greater than 1.80 for familiarity) compared with 2.31 for the 10 members in Group2 (which is also much greater than 1.59 for familiarity). Also, average standard deviation is 1.01, approximately twice as big as for the second group: 0.55. Turning back to the results from the previous section about familiarity we find these values very encouraging. Despite the fact that members in the first group know others preferences on average more than in the second group, they do not trust their peers music-wise as much as the members in the second group.

By observing the information displayed on each raw from Table II corresponding to each participant's rating we report that no participant in Group1 gave the same trust ratings to all other group members. In Group2, 3 participants, namely P14, P16 and P17 gave the same trust values ("3") to all their peers.

Table II. Trust values for Group1 (top) and Group2 (bottom)

Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9	Average	Std.
P1		2	2	3	2	2	2	2	2	**2.1**	0.35
P2	2		5	1	4	5	5	2	3	**3.4**	1.60
P3	3	5		3	4	1	1	3	1	**2.6**	1.51
P4	5	3	3		3	3	3	3	3	**3.3**	0.71
P5	1	1	1	1		2	4	3	2	**1.9**	1.13
P6	1	2	1	1	3		2	4	2	**2.0**	1.07
P7	1	1	1	1	3	1		1	2	**1.4**	0.74
P8	1	1	1	1	3	2	3		2	**1.8**	0.89
P9	1	1	1	1	1	1	4	1		**1.4**	1.06

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Average	Std.
P10		1	1	1	2	1	1	3	3	2	**1.7**	0.87
P11	3		3	3	3	3	3	3	3	3	**3.0**	0.00
P12	3	1		2	1	4	1	3	3	1	**2.1**	1.17
P13	2	1	4		1	4	1	3	3	2	**2.3**	1.22
P14	4	2	3	3		4	3	4	2	2	**3.0**	0.87
P15	1	1	1	1	1		1	1	1	1	**1.0**	0.00
P16	3	3	3	3	3	3		3	3	3	**3.0**	0.00
P17	3	3	3	3	3	3	3		3	3	**3.0**	0.00
P18	1	1	1	1	1	1	1	1		2	**1.1**	0.33
P19	4	4	2	2	2	2	2	4	4		**2.9**	1.05

4.3 Social Networks

The directed graphs included below (Figure 3) present the familiarity (up) and trust (down) networks of the two groups: "Cool party" to the left (P1 to P9) and "Legendary party" to the right (P10 to P19). The two graphs show that members positioned at the center of the graphs are mostly influenced by those who are at the exterior. These members are most familiar with their peers' music tastes most. Reciprocally, the participants positioned at the exterior of the trust graph are mostly trusted by others given a large number of weighted edges oriented towards them. The higher the weight given by participant Pi to Pj the thicker the directed line connecting them. We note from the upper-left graph that participants P1 and P4 are only familiar with P3 and P2's music preferences. Conversely, P2 stated that he is most familiar with the majority of other members' tastes. This phenomenon suggests that familiarity and trust information are not symmetrical. Very thin one-sided or two-sided oriented arrows are noticed for all 4 figures in green and red: they represent the smallest ratings which were submitted by our subjects which are all "1"s.

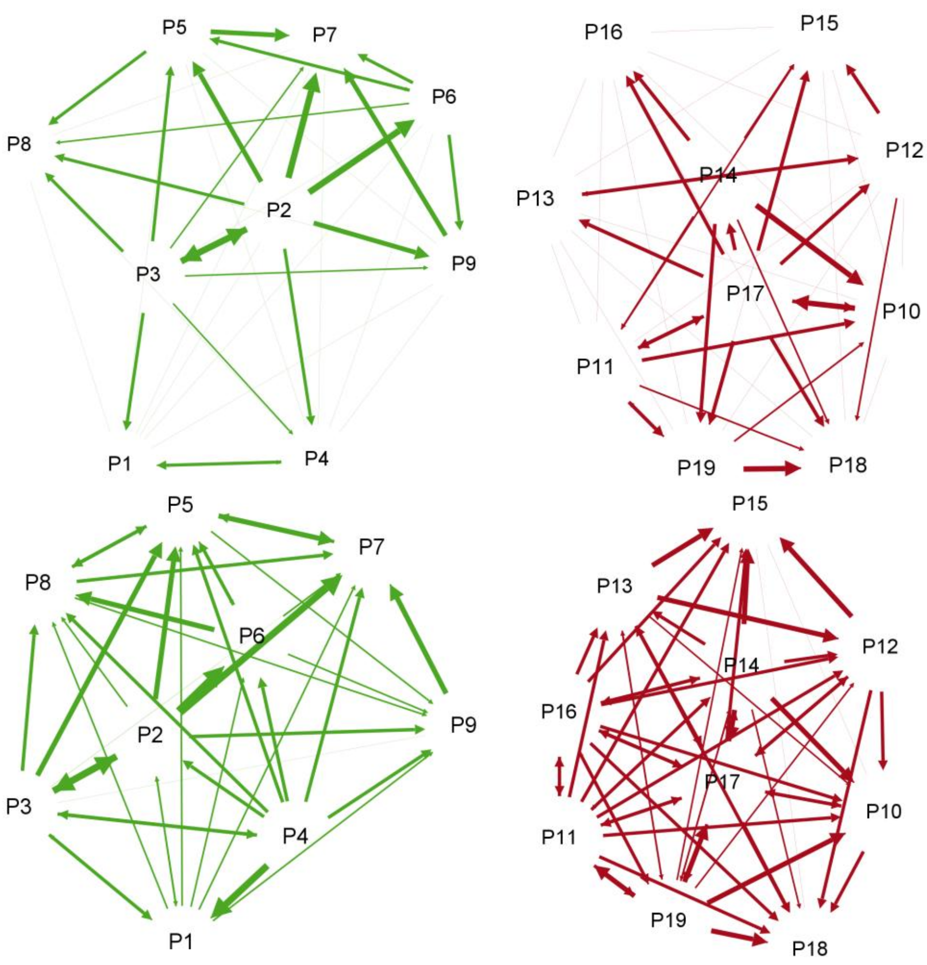


Figure 3. Familiarity and trust networks for the 2 groups

Comparing music familiarity among the two groups (upper graphs) we note tighter relationships in the second group: P2 and P3 in the first group take the role of P14 and P17 in the second one. Their central positioning shows that these subjects are the ones who are most familiar with others preferences (according to our familiarity and trust questions from the online questionnaires). Similarly, trust relationships are stronger in the second group given the higher edge-weight density. In this group we report a higher symmetry of members' familiarity whereas in the first group a sub-network formed by P2, P5, P6 and P7 can be observed.

Now, comparing the green and red networks among them-selves we remark that trust is determined by familiarity: the two upper graphs have sparser relationships than the bottom ones. We interpret this as follows: participants are more inclined to trust their peers' music tastes in their group even though they are not necessary familiar with such tastes. In a low impact domain such as music, the influence of others has been observed to be quite high, thus popularity plays an important role for individual songs or playlist recommendation.

The 4 graphs also allow deriving membership bounds which are useful for the study of reciprocal relationships among participants. For instance, the edge between P2 and P3 in the upper-left graph corresponds to the same score (5) given by both members reciprocally as a measure of their familiarity one to another. Similarly the edge between P17 and P18 corresponds to scores of 4 and, respectively 3, between the 2 members. The influential participants are those positioned at the exterior of the graph: most thick arrows are pointed towards them. Their music tastes are best known by others in the group.

It also worth noting that while for the second group (in red) P14 and P17 are positioned at the core of the graph for both familiarity and trust (these are the users who know and trust most all other members), for the first group (in green), P2 and P3 are users who are most familiar with others' music tastes while P2 and P6 trust all others the most. This represents a very important finding for our aim to understand group influence and group alignment. On the one hand we notice the high importance of trust for individual decision making and on the other the role played by familiarity in music discovery.

4.4 Familiarity and Trust Comparison

Out of the 19 participants only P2 submitted higher familiarity than trust ratings. Overall, participants are more open to discover their peers' music tastes in the second group, even though they are less familiar with these tastes which they trust a priori.

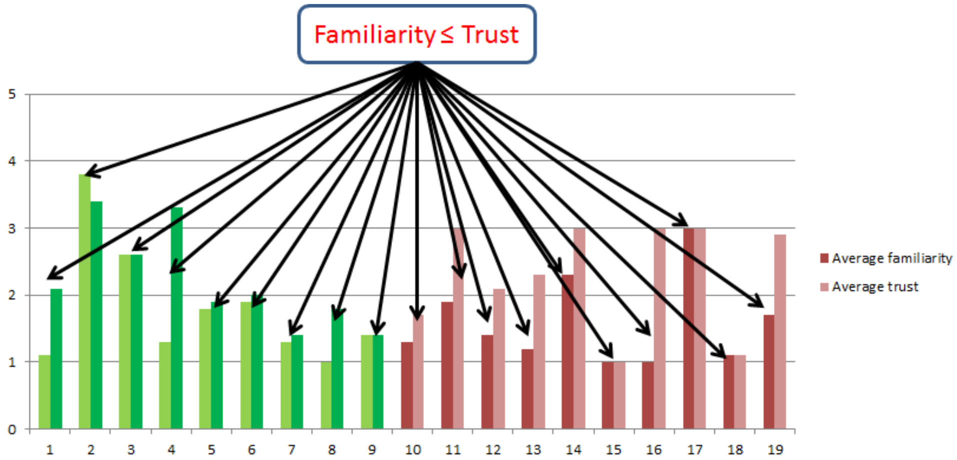


Figure 4. Average familiarity and average trust per participant

On the one hand, average familiarity across the two groups is similar (1.80 vs. 1.59 - Table III) whereas on the other, average trust is higher for the second group (2.31 vs. 2.21 – Table III): members trust others' more given higher connectivity of this group, i.e. higher familiarity and trust ratings to other members.

Table III. Average familiarity and trust for both groups

Group	Average Familiarity	Average Trust
1	1.80 (SD=0.81)	2.21 (SD=1.01)
2	1.59 (SD=0.55)	2.31 (SD=0.55)

In Figure 4 familiarity and trust ratings for the first 9 participants to the left is marked with light green and dark green, respectively, while the same information is marked with light red and dark red for participants 10 to 19 in the second group to the right.

One limitation of the above collected data of familiarity and trust is due to the fact that some participants submitted equal ratings (such as 1 or 3 for all others as marked in the previous 2 sections). Since some subjects used a balanced distribution (all ratings from 1 to 5) across group members and others relied only on few ratings it is difficult to produce a detailed comparison of familiarity and trust networks, respectively. An alternative rating solution would have been to consider ordered lists among participants submitted by each subject. In other words, subjects could have specified more balanced familiarity and trust distributions across their peers by giving, e.g. 2 ratings of "5", 2 ratings of "4", 2 ratings of "3", 2 ratings of "2" and 2 ratings of "1".

4.5 Familiarity and Trust Correlation

Another interesting phenomenon between familiarity and trust values is that they are highly correlated. For the second group we report a slightly higher correlation. With the exception of few closest members, some participants gave the same score to everyone else in the group, most times the minimum of 1. In some other cases few subjects chose the middle value of 3 for all members for which they could not differentiate familiarity and trust values.

Table IV. Correlation between familiarity and trust for the 2 groups

Group	Correlation value
1	0.613
2	0.619
1&2	0.587

In Figure 5 we include few examples of familiarity and trust rating behavior of 4 participants in Group1. The trust ratings of P1, P2, P3 and P4 are represented using a bar-plot-type chart. We use this visual representation to highlight the rating styles' differences among participants: P2 gave 3 ratings of "5" whereas P1 only 1 rating of "3" and all others "2". Similarly, P4 gave only 1 rating of "5" to P1 and all other members the same rating of "3". P3 devised differently other members by trust: one member (P2) who he trusts the most (rating of "5"), 1 member who he trusts to a large extent (P5) and 3 other members that he somehow trusts (P1, P4 and P8).

Figure 5. Example of familiarity and trust assessment

5. GroupFun Experiment

In addition to the online questionnaire, all 19 subjects participated in the GroupFun phase of the experiment which lasted approximately 3 hours per group.

5.1 The GroupFun System

GroupFun is a web application that helps a group of friends to agree on a common music playlist for a given event they will attend, e.g. a graduation ceremony. First, it is implemented as a Facebook plugin connecting users to their friends. Secondly, it is a music application that allows individuals to manage and share their favorite music with groups they create or join. In GroupFun users can listen to their own collection of songs as well as their friends' music. With the collective music database, the application integrates friends' music tastes and recommends a common playlist to them. Therefore, the application aims at satisfying music tastes of the whole group by aggregating their individual preferences.

Our design is based on previous user studies and pilot tests regarding group decision and interaction aspects. GroupFun's structure includes 3 sub-pages: "Home", "My Groups" and "My Music". In the first one, users see 4 playlists: one containing most popular songs, one used at a previous event, another one including recent uploads and the last one from a group party. They can listen to each song in each of the playlists. In the "My Groups" page users create groups, upload and rate their music, invite friends and hear the group's songs (Figure 6). Finally, in the "My Music" page users see their contribution to GroupFun: for each song the interface displays the associated group, the user rating and its name and artist. Users can also listen to their individual uploads using a wheel-like play button.

One of the most important characteristics of GroupFun is that it combines music, friends and groups together. This distinguishes the system from related work. First, it meets users' expectations for conveniently organizing their individual music. Then, it supports effective communication among friends through ratings. Thus, users participate in social activities while enjoying their common collection of music.

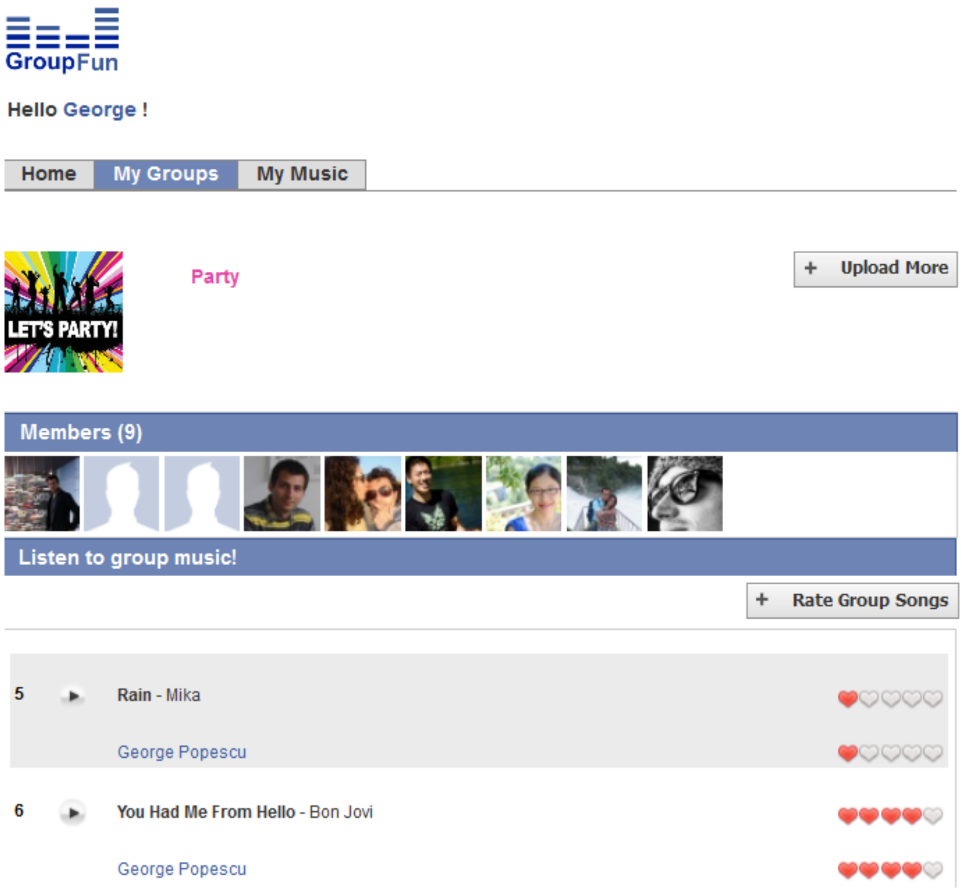


Figure 6. Group interface in the GroupFun recommender system

GroupFun is both a group music recommender system and a Facebook application. In order to access the system participants needed a Facebook account. Most participants used their personal account as we guaranteed that the application will not post any information on their wall and collected only username, name and family name information. The other participants created a temporal account which they erased after the experiment. These participants were not Facebook fans and did not want to use the social website for their personal life. They used random names on their basic profile which had no photos or background information.

As mentioned before, users connected to the system to create 2 groups: “Cool party” and “Legendary party”. Each subject was assigned the main task of uploading several songs and rating as many as desired group songs. After uploading and rating self-songs each saw a list of songs already uploaded to the group by other members. They listened and rated group songs individually, i.e. without any authoring information – the interface displayed only the name of the song and the artist without the group member who contributed to the song.

The individual rating interface used for participants to elicit their preferences is included in Figure 7. Each song is indexed in the left hand side according to the upload time. A multi-functional play button allows users to play, pause and fast-forward the selected song. In the center the song's name is included in bold followed by the artist's name. Finally, members rate group songs by clicking on the appropriate heart numbers corresponding to a rating ranging from "Strongly dislike it" (rating of 1) to "Strongly like it" (rating of 5). The interface also includes GroupFun's logo (at the top), the 3 tabs of the application, the group's name and a box displaying all group members.

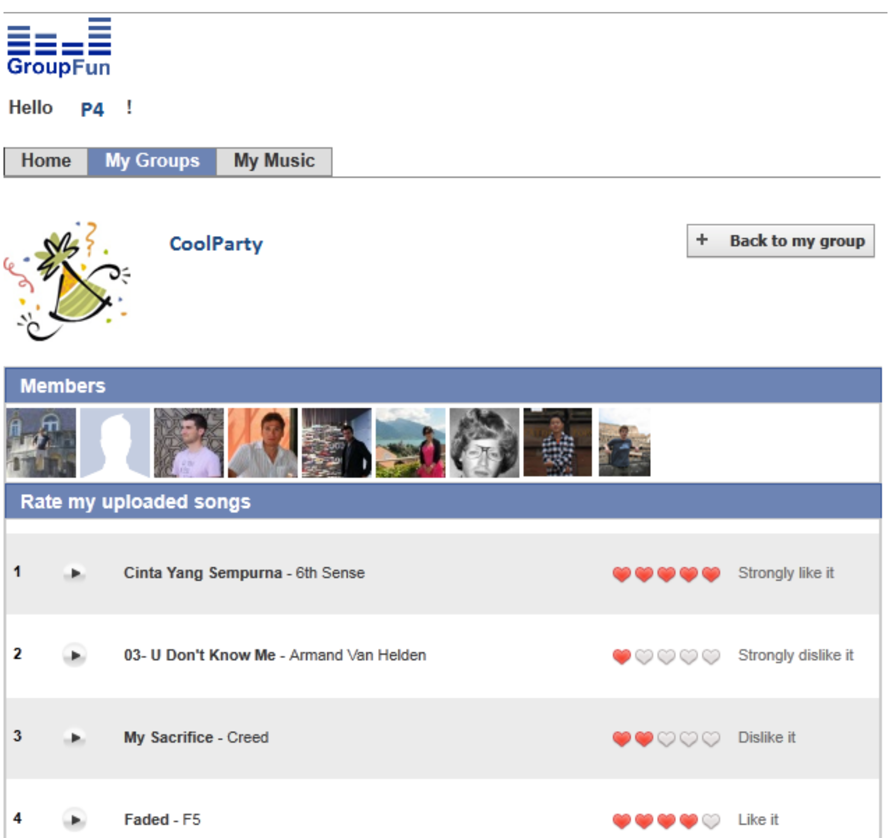


Figure 7. Individual rating interface in GroupFun

5.2 Contributed Group Songs

Members of the “Cool party” group contributed with 39 songs (2h and 51m listening time, 212 MB disk size) while those of the “Legendary party” with 47 songs (2h and 55m listening time, 202 MB disk size) all in .mp3 format: an average of more than 4 songs per user. The total library collection size was 414 MB. All contributed songs are listed in the appendix and include the title, artist and YouTube URL that was used to download the music protecting us from copy-right aspects.

5.3 Contributed Group Ratings

In the first group the 9 participants submitted a total of 203 ratings – most ratings submitted by P4 (39 ratings) and least by P2 and P6 (9 ratings) - yielding a group satisfaction score of 3.27 (standard deviation SD=0.43) – most satisfied member is P2 (average satisfaction 3.9) and least satisfied P7 (average satisfaction 2.7) computed as the average among all individual preference ratings.

Table V. GroupFun statistics: Group1 (top) and Group2 (bottom)

	P1	P2	P3	P4	P5	P6	P7	P8	P9	#
#ratings	20	9	19	39	26	9	37	29	16	**203**
Average (39)	3.3	3.9	3.6	3.2	3.7	3.4	2.5	2.7	3.1	**3.27**
Average (24)	3.3	4.0	3.6	3.0	3.8	3.5	2.1	2.2	3.2	**3.12**

	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	#
#ratings	10	13	21	16	10	16	13	23	33	41	**196**
Average (47)	4.2	3.7	4.2	4.4	3.7	3.4	3.2	3	3.3	3	**3.61**
Average (24)	4.1	3.6	4.2	4.3	3.6	3.7	3.2	2.5	3.4	2.9	**3.52**

In the second group the 10 participants submitted a total of 196 ratings for all 47 songs yielding a higher group satisfaction index of 3.61 (SD=0.47). In this group most satisfied members are P10, P12 and P13 whereas least satisfied are P17 and P19.

Now, we restrict the same computation to only the (first) 24 songs that are used for the 3rd phase of the experiment. In this case, the average values decrease slightly. This is because in the selection of 24 songs few songs contributed by each participant were eliminated. Those songs contained highest ratings as they were uploaded by their "owners". The order remains unchanged: Group1's satisfaction decreased from 3.27 to 3.12 while Group2's satisfaction decreased from 3.61 to 3.52. Thus, the second group is the most satisfied one for both conditions.

5.4 Top 10 Highest Rated Songs

We include a selection of Top 10 songs (out of 39 and, respectively 47; Table VI) rated by at least 2 members (in order to avoid dictatorship) for both groups taking as metric the average score without any limitation regarding the number of users rating that song. We thus eliminate the songs having a maximum rating from one single user. We observe a higher homogeneity for songs' ratings in the second group. While in the first group there are few great songs that would please everyone, in the second group the dispersion of ratings is much smaller: it becomes more difficult to cluster good songs from undesired ones. However, such ratings must be weighted by the number of users who gave a high rating, those who gave a low rating and those who did not rate at all.

5.5 Top 24 Most Rated Songs

Out of the 39 and 47 songs from the GroupFun phase we selected top 24 songs for the eye-tracking phase of the experiment based on the number of ratings. Each of the 24 songs got more ratings than the remaining 15 for Group1 and the remaining 23 for Group2, respectively. The main idea behind this selection was to have the greatest number of ratings to compare for the individual preference elicitation and the group-influence preference elicitation phases. Having a larger dataset of ratings would allow us to understand both individual decision behavior and group dynamics.

We aim to compare initial ratings with re-ratings to see if one individual changed his decision given the preferences of others. However, since not all subjects rated the 24 songs in the first phase we expect the change rate to vary according to the number of re-rated songs. It is easier for someone to remember one rating for a specific song in GroupFun and to re-state that preference rather than having to re-rate 24 songs one more time and remember the ratings for all of them. From the rating point of view we are interested in quantifying the score differences and develop our social influence argument based on the eye-tracking fixation points times and eye-gaze correspondences data. We analyze each submitted rating in relation with the information included in the displayed eye-tracking interface.

In the appendix section there are two tables containing all the 24 songs selected per group together with the initial poll of all songs in the library.

Table VI. Top 10 songs in Group1 (top) and Group2 (bottom)

Nr.	Track	Artist	Average
1	I Know You Want Me	Pitbull	5
2	Meaw	Deadmau5	5
3	Duel Of The Fates	The London Symphony Orchestra	4.8
4	Basket Case	Green Day	4.6
5	Can't Take My Eyes Off You	Frankie Valli	4.2
6	Ya Basta!	Gingala	4
7	Lost Together	Sophie Sugar vs. Sunlounger feat. Zara	3.7
8	System Of A Down	Chop Suey	3.6
9	Crave	Renaud	3.5
10	Stirb Nicht Vor Mir	Rammstein	3.3

Nr.	Track	Artist	Average
1	Life Is Wonderful	Jason Mraz	4.4
2	Baazi	Siavash Ghomayshi	4.3
3	This Is The Life!	Amy Macdonald	4.3
4	Basket Case	Green Day	4.2
5	I'll Be Your Mirror	The Velvet Underground	4.2
6	Wake Me Up When September Ends	Green Day	4.1
7	Someone Like You	Adele	4
8	Don't Worry Be Happy	Bob Marley	4
9	Poker Face	Lady GaGa	4
10	It Will Rain	Bruno Mars	4

6. Eye-Tracking Experiment

We designed our eye-tracking experiment aiming at understanding how people perceive other members' ratings. Furthermore, we were interested in analyzing the type of correspondences they make through visual contact. Our eye-tracking interfaces were restricted to include only the minimum relevant song, names and ratings information.

6.1 Experiment Steps

The eye-tracking study lasted for 25-30 minutes. During the gaze-recording experiment each participant participated in:

Step1. The experiment's admin first debriefs each participant on the nature of the experiment explaining the main process flow consisting in the 3 sub-steps presented in Figure 1 – Experiment design (red box).

Step2. The admin assists each participant during the calibration process of the eye tracking system. He then saves the calibration data for each user and loads it at the beginning of the recording session.

Step3. The user fixates the iPod headsets in his ears and positions him-self in a comfortable position in front of the eye-tracking monitor device. He uses his right hand to locate the "forward" and "pause" buttons on the iPod which are useful to shift to the next song (forward button one-time press), fast-forward the current one (forward button kept pressed), pause the song when ready (pause button one-time pressed) or continue to listen to the current song (pause button one-time pressed again).

Step4. At this moment the most important part of the experiment starts. Each user listens to one song and sees a corresponding interface at a time. In each interface are presented the title and artist information (at the top), a list of other members' names (to the left) and a list of their ratings (to the right). There are 24 recommended songs on the iPod corresponding to the 24 interfaces customized for each user. Since each participant sees only the ratings of others and not those of him-self we modified each interface for all corresponding conditions so that the current participant would be presented the relevant social information adapted to him. When ready with the song evaluation the participant gives a rating from “1” to “5” to the song he just listened to. The participant proceeds to the next song by clicking on the “forward” button and any keyboard key on the computer controlling the eye-tracker.

Step5. To conclude the study, the admin collects general comments through open discussions with each participant in order to assess all users’ overall perception of the experiment.

6.2 Experiment Progress

A Tobii 1750 eye-tracking device was used to track user attention when facing the set of 24 interfaces displayed to them. In this phase of the experiment the main user task was to rate 24 songs while shown the 24 corresponding interfaces. We asked our participants to listen to 24 songs using an iPod shuffle for as long as desired to make a rating decision. In the same time we displayed an interface which included the song’s name and artist at the top and a group of other members’ names together with their ratings at the bottom. We controlled both the rating values and the persons’ names in the interface according to the conditions which will be explained in the next sections. The ratings were not the real ones from the GroupFun phase of the experiment but were generated according to some popularity and trust conditions explained in the following.

Before each experiment session we informed each participant of the experiment’s objectives and explained the general labeling on each of the interfaces. They were told that their eye-gaze would be recorded and were suggested to find a comfortable position in front of the eye-tracker so that the monitor would output a clear recording of their gaze. At the beginning of the experiment each participant had a pair of headsets playing the songs according to the interface displayed to them. When deciding on the rating values subjects pressed the pause button on the iPod and submit their rating for that song. With a key press they would move on to the next song and interface displayed on the eye-tracker. Each subject experienced the same conditions structured in the same order regarding when he performed eye-tracking study (e.g. the 2nd and the 9th participant in the first group saw the same experiment conditions in the same order).

Subjects could listen to music and gaze at the interfaces for as long as they felt it was needed to decide. As a consequence, we expected that some participants would submit their rating very fast noticing that the current song being played is one from their contribution: this was the case for 2-3 songs per participant. As such they would most probably re-submit the highest rating of “5” to that specific group contribution they made. We noticed that in the music context people tend to be more critical about others preferences for which they give smaller ratings, rather their own.

Similarly, when hearing one song that they would not like at all they would rapidly stop the song and give the smallest rating of “1”. In between these extreme cases, which occurred rather rarely, subjects took the time to look attentively at the interfaces and produce a preference elicitation eye-gaze-based judgment.

The eye-tracking interface experiment flow is summarized in Figure 8. Each set of 3 conditions highlighted in each box includes two songs. The order of the conditions displayed in the experiment is marked with the blue arrows As an example, the first condition (top-left-most orange box) includes 3 factors: “other 8/9 members” (the interface displays the ratings of all other group peers), “consensus” (all ratings have small standard deviation) and “popularity” (the order of the items in the interface is given by the ratings column - highest ratings of “5” at the top). In the same fashion, the last condition (bottom-right-most blue box) shows the ratings of “2 closest members” who have divergent ratings and the 2 are sorted by trust: most trusted member at the top.

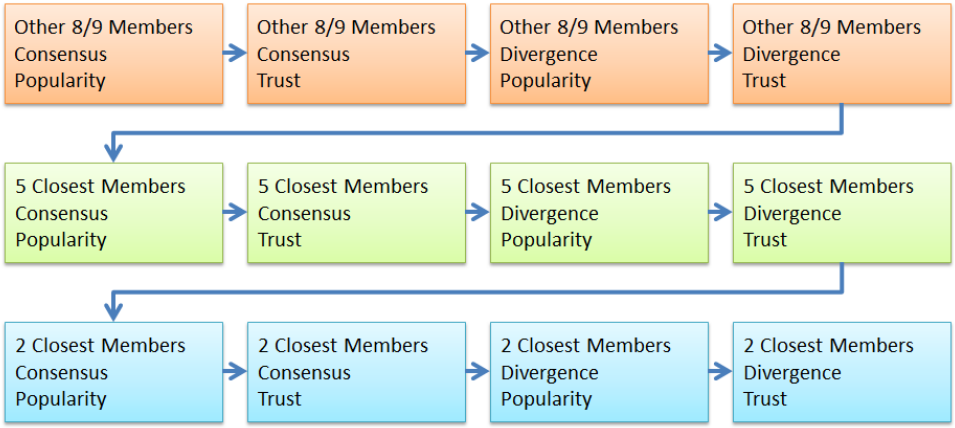


Figure 8. Eye-tracking experiment interface design flow

6.3 Interfaces

There are several essential observations we need to make about our experiment. First, it is tailored for each user considering personalized interfaces for each participant. Thus, we eliminated the current participant's name from the list and updated all other interfaces with his other peers' names and developed the appropriate conditions given the trust and ratings values.

Separately, we considered 8 interfaces corresponding to all remaining group members, 8 with closest 5 members and the last group of 8 interfaces with 2 closest members. The terminology "closest members" is used to denote the highest trust values extracted from the questionnaire. Each of the last 16 interfaces was adjusted by the trust values and not by familiarity or combined values. We assessed that, in general, trust ratings were more reliable in assessing the closeness between group members. The submitted trust ratings also covered a wider range compared with familiarity values allowing us to identify the 5 closest members and 2 closest members better. In particular we could distinguish those closest members from the rest better with trust than with familiarity. Not only that trust ratings were higher than familiarity ratings but subjects rated others considering a more sparse distribution which allows for the distinction between 5 most trusted and 2 most trusted members making our selection more accurate.

The design of each interface is based on the fact that participants, as decision makers undertaking the rating task, use the social information displayed to them to adjust their ratings to those of the group, result which is sustained by the social conformity theory. Therefore, we aim at understanding the individual decision making process together with the main contributing factors.

In specifying a rating for a song that he listens to, a user explores the interface presented to him with the design shown in Figure 9. The eye-tracking device records his eye movements and fixation times while browsing throughout the interface. Through our design we expect our participants to create numerous horizontal correspondences between the name's area delimited with blue color and the ratings area displaying the rating and marked with red. These correspondences are essential for understanding which other members contributed most to one's preference decision. As the number of names and ratings becomes smaller participants create more dense correspondences.

In order to have clear regions on the interfaces for eye-tracking data collection we divided all interfaces into 3 areas of interest (AOIs, Figure 9):

(1) AOI Song = AOIS: the grey rectangle at the top;
(2) AOI Users = AOIU: the left part containing participants' names;
(3) AOI Ratings = AOIR: the right part containing user ratings.

The interface from Figure 9 contains a random list of first names and family names replacing real participants' names. This figure was generated with the purpose of showing the reader a clear interface containing the information which our participants saw on their screens. Thus each subject in the first group saw first 8 interfaces of 8 other members in the group whereas each participant in the second group saw first 8 interfaces with 9 other members and their ratings. The next 16 interfaces for both groups contained the same amount of information: 8 for 5 closest members and 8 for 2 closest members. This particular figure corresponds to the following conditions: all other members (9 other members in Group2), divergence (highest standard deviation across ratings ranging from "1" to "5") and trust (the ordering is done by the names column to the left).

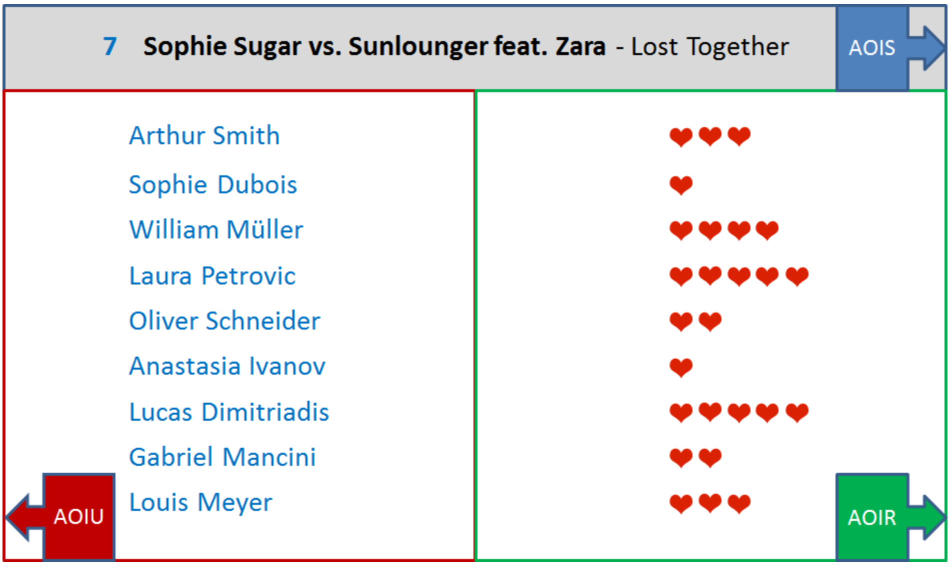


Figure 9. Example of eye-tracking type interfaces

Altogether, the 24 interfaces were conceptualized as follows:

(1) First group of 8 interfaces: all other group members (Figure 8, orange top row);

(2) Second group of 8 interfaces: 5 closest members (Figure 8, middle green row);

(3) Third group of 8 interfaces: 2 closest members (Figure 8, blue bottom row).

Furthermore, we split each of the 8 into:

(1) Controversial: highest standard deviation of ratings (Figure 8, divergence);

(2) Non-controversial: lowest standard deviation of ratings (Figure 8, consensus).

Every next set of 2 consecutive interfaces corresponds to one of the additional ordering condition:

(1) Trust: the ordering of names shows the most trusted members at the top and least trusted at the bottom;

(2) Popularity: songs are sorted by their received rating and are information is displayed from highest ratings at the top to lowest ratings at the bottom.

A summary of all of the above conditions is visually presented in Figure 10. For each box we displayed 2 songs consecutively. Thus, the same set of conditions would benefit from 2 ratings and 2 eye-gaze data sets. The flow follows the same rule as in Figure 8: from top-left to bottom-right in zigzag.

The 3 rows in Figure 10 display all other group members (8/9 for Group1 and Group2, respectively), 5 and 2 closest members based on trust values submitted by each participant with respect to other members. The 4 columns in the table group 2 types of opposing conditions: consensus vs. divergence and trust vs. popularity. The first group refers to the fact that the participants in the group agree with a certain item/song having a small standard deviation of ratings (consensus) whereas the opposing condition present ratings with high standard deviation (divergence). Trust and popularity refer to the order in which the information is displayed on the screen from top to bottom. In the "trust" condition, the sorting is done with respect to the participants names (from AOI Users): at the top are placed most trusted participants whereas at the bottom the others. Similarly, in the "popularity" condition participants and their ratings are sorted by the ratings column (from AOI Ratings) to the right: highest rating at the top (5 hearts) and lowest at the bottom (1 heart).

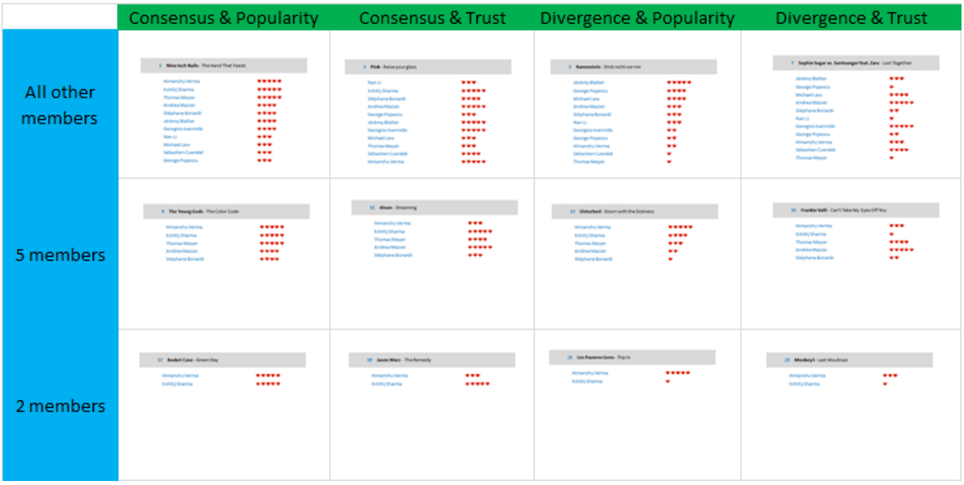


Figure 10. Design examples for all 24 eye-tracking interfaces

As an example, each participant would first see one interface displaying Song1 in which all other members agree on similar ratings (consensus or lowest standard deviation of ratings) and songs are sorted by ratings: from highest to lowest values. In the second interface he would see Song2 in the same condition. In interface three he would see Song3 together with all other members' ratings sorted by trust: most trusted members at the top and least trusted at the bottom from highest to lowest rating. Song4 would be in the same condition. In interface 5 the current subject would see divergent ratings of all other members sorted by the songs rating. Song6 would display the same information Song7 is included in an interface in which all other members submitted divergent ratings and their names are sorted by trust. The same condition applies for Song8. From Song9 until Song16 only the names and ratings of 5 most trusted members are displayed. Finally from Song 17 to Song 24 only the names and ratings of 2 closest members are included.

6.4 Eye-gaze Output

Eye-tracking technology allows to output different types of visual representation in the form of: eye-gaze, fixation points, circles, bubbles, gaze maps, etc.

We preferred the fixation-point representation (example included in Figure 11) as the most useful for extracting horizontal correspondences and fixation times. Below the participants' names from AOI Users in the left hand-side of the image were erased for the purpose of preserving the privacy of our subjects. The blue circles represent the main fixation points that the eye-tracker detected, the blue lines represent the eye's movements from one element to the next and the numbers inside the circles represent the number of milliseconds for each fixation.

This graphic schema is a powerful representation for: (1) extracting fixation times for each of the areas of interest based on which we detect users that are most influential for each participant by summing up the fixation times per interface and (2) analyzing fixation directions: (nearly) parallel lines (between AOI Users and AOI Ratings) correspond to the fact that participants associate other members with their ratings whereas vertical lines (within AOIs) denote the fact that users are in search of either user or rating information which is not coupled: user-rating.

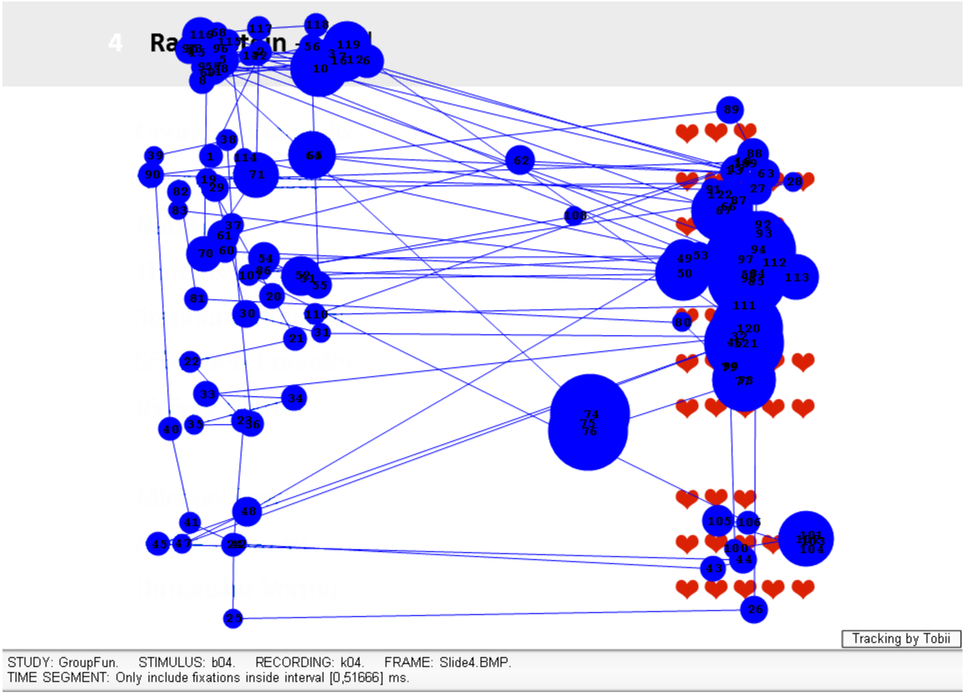


Figure 11. Eye-tracking output using fixation-point map

Considering the output provided in Figure 11 we make the following observations. First of all, the current participant produced numerous gaze correspondences, both horizontal and vertical. This suggests that his decision is strongly influenced by the information included in the interface. Secondly, the fixation times are rather large especially for the ratings column to the left (upper part) denoting the user's interest was focused in this area. The participant also skimmed more quickly the left column. This effect was expected: the subject already knows other members' names, but not their ratings. It becomes natural for our subjects to spend slightly more time in the ratings area of interest.

6.5 Areas of Interest and Time in AOI

An extensive dataset containing fixation points (measured in seconds) was collected based on 19 participants x 24 interfaces = 456 eye-tracking output images. We recorded the total time corresponding to each user-name and rating for each of the three AOIs and for each participant. A summary is presented in Table VII with the following abbreviations:

- TTI = Total Time Indicated;
- TTC = Total Time Computed;
- AOIS = Area Of Interest Songs;
- AOIU = Area Of Interest Users;
- AOIR = Area Of Interest Ratings.

The total time indicated for Group1 is 4'867.5 seconds and for Group2 6'436.8 seconds. In total the 19 participants gazed for 11'304.3 seconds, which means 595 second per participant, on average, i.e. around 10 minutes of continuous looking at the interface, without breaks. The continuous gaze is 9'417.1 seconds for both groups, i.e. 495.6 seconds per participant or 8.26 continuous gaze-recording per subject.

The proportion between the "Total Time Computed" (TTC) and the "Total Time Indicated" (TTI) gives the eye-tracker's efficiency for each group: 84% (4080.7/4867.5) for Group1 and 83% (5336.4/6436.8) for Group2. This means that from the total time spend in front of all interfaces 84% of eye-gaze information was collected for the first 9 participants and 83% of all eye-gaze was recorded for the last 10 subjects. The TTC is further divided by the 3 conditions corresponding to interface design: “8/9” other group members, 5 closest members and 2 closest members.

Table VII. Time view-statistics (in seconds) per group

Group	Total Time Indicated	Total Time Computed	Time in AOIS	Time in AOIU	Time in AOIR
1	4867.5	4080.7	1536.7	1382.2	1161.8
	8 other members	1642.1	386.6	640	615.5
	5 closest members	1284.7	543.8	406.9	334
	2 closest members	1153.9	606.3	335.3	212.3
2	6436.8	5336.4	1517.4	2188.9	1630.1
	9 other members	1793.5	274.6	891.7	627.2
	5 closest members	1872.4	464.7	786.5	621.2
	2 closest members	1670.5	778.1	510.7	381.7

Several preliminary conclusions yield from table VII:

(1) Overall, for all interfaces, participants spent a different amount of time in all 3 AOIs.

(2) For all members, users of both groups looked mostly at the AOIU, then at their ratings and finally at the songs name.

(3) For 5 closest members the situation is different per group. The first group looked mostly at AOIS and least at AOIR. Since the songs’ title and artist does not contain any social influence power but only the horizontal correspondences between AOI Users and AOI Ratings, users were least influenced. For the second group users looked mostly at AOIU, then at their ratings and finally at AOIS being mostly influenced.

(4) For 2 closest members, participants from both groups looked most at the song’s name and artist, then at other members and finally at their ratings. This suggests that they were least influenced. However, since the interface area was reduced in size it is somehow natural that they looked less than before at these last two AOIs.

The total time computed is higher for the second group compared with the first one. The relative difference of about 80 seconds on average per participant (e.g. 5336.4 / 10 = 533.64 seconds / member in Group2 vs. 4080.7 / 9 = 453.41 seconds / member in

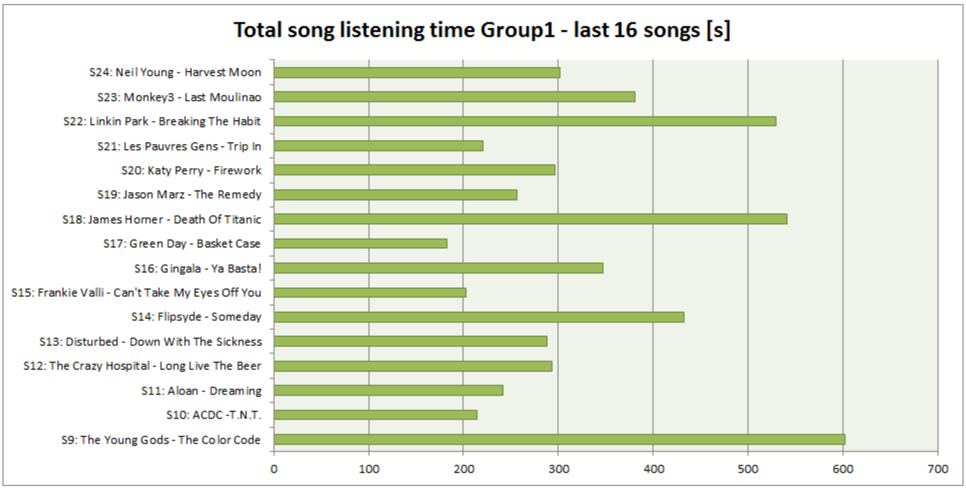
Group1) is not only due to the one extra group member, but to the overall interest of this group's members to see others' tastes.

A more in-depth analysis regarding visualization times is performed in the results section. Relative percentages for each AOI are explained together with in-depth analysis and detailed explanations regarding the familiarity and trust networks.

6.6 Song Visualization Time

In addition to the above, each song received a different sight attention from all members. We include in the next 2 figures a selection of the last 16 songs corresponding to 5 closest and 2 closest members because these 2 conditions show larger differences per group. The average listening time per group for the first 8 interfaces does not demonstrate large watching times differences as all members had about the same behavior and spent about the same time scanning all the information displayed to them.

The respective watching times for Group1 for the last 16 songs are listed in Figure 12. We report that some songs, such as S9 (602 seconds), S18 (541 seconds) and S22 (530 seconds), have as much as 3 times more watching time than others, such as S17 (only 183 seconds). Given that the number of displayed users for all 8 songs was 2, we formulate the following possible explanations: either the users did not know the respective songs and discovered it while listening or they enjoyed it more than others and wanted to continue listening.



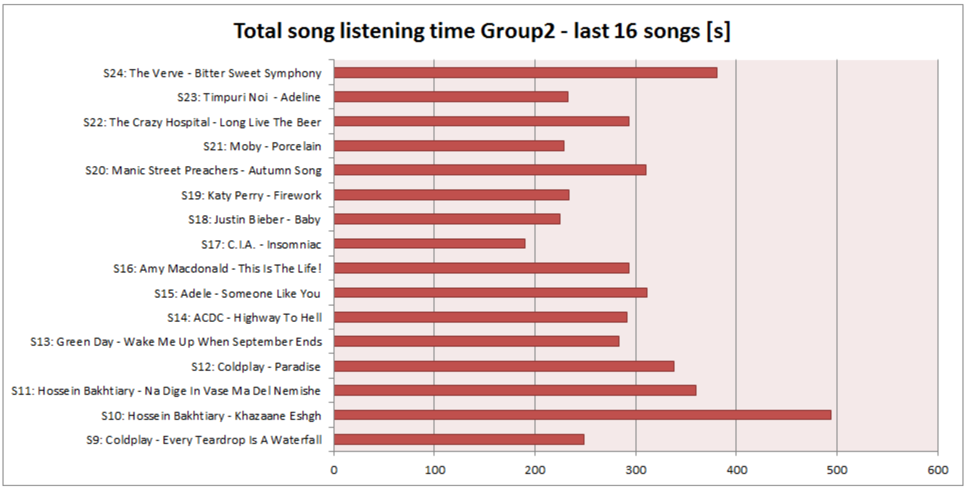


Figure 12. Total listening time for all members: Group1 (top) and Group2 (bottom)

In Group2 the situation is significantly different: all total listening times for all 16 songs is concentrated closer to the average and clearly shows smaller dispersion. With the exception of S10 (which is an Arab song) and had 494 seconds total time, all other songs have a listening time between 381 seconds (S24) and 190 (S17).

Other factors such as song's language and music genre can contribute to one's rapid acceptance or rejection of a song. As a general trend, popular songs tend to have longer listening times, whereas serendipitous contributions tend to have shorter listening times. All visualization times per song and per group are included in the Appendix section at the end of the report clearly marking this general phenomenon. Additionally, it is interesting to notice that certain songs received particular attention despite their reduced popularity due to the effect of music discovery.

Another dimension that we investigated was to test weather users would look most at their own contributed songs or others' songs. The pattern shows that there is no correlation between the two. Rather than this, users spent more time for the interfaces displaying to them only 5 or 2 closest members together with their ratings. The song itself is not an important factor for visualization.

7. Results

By comparing the two groups we prove that the network structure is strongly correlated with the eye-tracking data. We extract eye-tracking fixation points associated to duration times and produce a qualitative analysis of participants' gaze.

Moreover, we compare individual satisfaction for the GroupFun system with the individual satisfaction through the eye-tracking group study focusing on the same songs users saw in both scenarios.

We start by relying on the statistical data derived from the collected data. Next, we present the eye-gaze patterns and summarize the main findings related to the relative duration time in each area of interest. With respect to the submitted ratings for each of the 24 songs per group we compute the preference correlation between each 2 participants and identify closest subjects in terms of music preference. In the "ratings change" subsection, one of the most important of this chapter, we include significant results which prove group alignment. First, we extract the ratings from the individual and group interfaces. Based on these contrasting values we compute the relative change rate per participant and per group determining a group influence score. We enhance our results by eliminating own songs and ratings. Finally, we analyze social influence with respect to group size on three layers: "all other members", "5 closest members" and "2 closest members". In addition to this, we compute change rate values for consensus vs. divergence.

7.1 Statistical analysis

We employ statistical tools to: (1) test the validity of our experiment with respect to our participants' rating behavior and (2) compare sets of 2 experimental conditions as defined through our study design.

7.1.1 Consistent Rating Behavior

We used the Chi-Square statistical test which yields whether actual data differ from a random distribution comparing the counts of categorical responses for the "8/9", "5" and

"2" closest members grouped in the 3 independent AOIs: AOIS, AOIU and AOIR. We obtained very low p-values for both groups and 3 conditions: p_AOIS_1 = 1.365E-11, p_AOIU_1 = 1.181E-24, p_AOIR_1 = 1.085E-48, p_AOIS_2 = 2.667E-34 and p_AOIR_2 = 0.001. Thus, the probability that participants rated randomly is almost 0!

7.1.2 Pairs of 2 Conditions Comparison

We performed a random effects one-tailed t test over beta values for each regressor across all subjects. The p-values smaller than a threshold of 0.05 are marked in green in the tables below.

In both groups we obtained higher p-values as the number of group members in the interfaces is reduced over all fixed conditions (Table VIII summarizes the main conditions): consensus, divergence, trust and popularity.

These results prove the statistical significance of our data as follows: we reject the null hypothesis that the two groups have the same average rating overall, for the first set of 8 interfaces and the second set, but not for the third one. Also when the divergence and trust conditions were applied the p-value is less or close to the significance level of 0.05 (Table VIII).

Table VIII. p-values comparing all conditions for both groups separately

Group1	**Fixed Divergence**	**8**	**5**	**2**	**Fixed Consensus**	**8**	**5**	**2**
	Trust vs. Pop.	0.01	0.114	0.887	**Trust vs. Pop.**	0.7	0.444	0.807
	Fixed Trust	**8**	**5**	**2**	**Fixed Popularity**	**8**	**5**	**2**
	Div. vs. Cons.	0.038	0.61	0.469	**Div. vs. Cons.**	0.573	0.74	0.474
Group2	**Fixed Divergence**	**9**	**5**	**2**	**Fixed Consensus**	**9**	**5**	**2**
	Trust vs. Pop.	0.072	0.298	0.45	**Trust vs. Pop.**	0.048	0.452	0.028
	Fixed Trust	**9**	**5**	**2**	**Fixed Popularity**	**9**	**5**	**2**
	Div. vs. Cons.	0.001	0.193	0.54	**Div. vs. Cons.**	0.472	0.247	0.417

In Group1 when the divergence condition is fixed participants did not differentiate between the order given by trust or popularity for the “8 other members” condition (p=0.01). Also, when the trust condition was fixed, they did not differentiate between consensus and divergence for the first set of 8 interfaces (p=0.038). The data shows that for “5” and “2” closest members the two groups differentiated among them-selves with respect to the ratings they submitted. This is a very important result showing that the 5 and 2 closest members have a strong impact on the submitted ratings. No p-values smaller than 0.05 were computed for the fixed consensus and fixed popularity in Group1.

In Group2 we obtained the same situation between consensus and divergence when the trust ordering was applied (p=0.001). Also a small p-value (p=0.072) was computed for the “9 other members in the fixed divergence condition. However, by opposition to the first group, when the consensus condition was fixed (the interface displayed ratings with smallest standard deviation) participants did not distinguish between the trust and popularity ordering for “9 other members” and “2 closest members”. For all other conditions, we obtain high p-values. Most notably, when the fixed popularity condition was applied (songs are ordered by ratings), participants did not differentiate between divergence and consensus.

From the data presented in Table VIII differentiating between the two groups we report that individuals generally do not distinguish between the ordering by trust or popularity for both groups (p=0.01, p=0.038, p=0.072 and p=0.01). Moreover, in the second group, when facing the 9 members and 2 members conditions participants do not distinguish between the ordering by trust or popularity (p=0.048 and p=0.028).

Statistical measures from Table IX allow us to formulate more general conclusions. First, participants were not sensitive to the popularity ordering overall (all p-values on the popularity row are large: 0.549, 0.543, 0.127 and 0.321). Secondly, when seeing all other (“8”/”9”) group members and their ratings (first set of 8 interfaces) subjects did not opt for neither of consensus or divergence. Rather than this, they gave individual ratings following own valuations. Finally, overall, the two groups have similar ratings, results are alike for the first and second set of interfaces (p=0.053 and p=0.046) and different for the

“2 closest members” condition (p=0.833). This is an expected outcome since the second group is more socially connected than the first one.

Table IX. p-values comparing al conditions in a between-groups analysis

Overall	All interfaces	8/9 members	5 members	2 members
Gr1 vs. Gr2	0.018	0.053	0.046	0.833
Consensus	**All interfaces**	**8/9 members**	**5 members**	**2 members**
Gr1 vs. Gr2	0.403	0.140	0.602	0.653
Divergence	**All interfaces**	**8/9 members**	**5 members**	**2 members**
Gr1 vs. Gr2	0.010	0.200	0.015	0.431
Trust	**All interfaces**	**8/9 members**	**5 members**	**2 members**
Gr1 vs. Gr2	0.003	0.040	0.176	0.126
Popularity	**All interfaces**	**8/9 members**	**5 members**	**2 members**
Gr1 vs. Gr2	0.549	0.543	0.137	0.321

For all 24 interfaces comparing Group1 with Group2, participants mostly diverged from the consensus and popularity conditions presented in the interfaces and gave similar ratings when the trust and divergence conditions were applied. This yields that individuals are more alike in both groups when shown divergent ratings. For such interfaces, they impose their own valuation. When shown most trusted members they align to the group decision specific for their group. Conversely, the groups differ most when their members face consensus, in the second group participants aligned more to the group output than in the first group, and popularity, the ordering of ratings is again more important in the second group.

From Table IX we report that individuals align more with the group decision when the group is smaller and deviate most from the group decision when the group size is biggest. Furthermore, when the divergence condition is applied for the last group of interfaces displaying 5 members we note that members in the second group align more to others preferences than in Group1 (p=0.015).

To conclude, the p-values included in Tables VIII and IX above are very useful to understand the conditions for which participants differentiate the most. We distinguish between group size, consensus vs. divergence and trust vs. popularity.

7.2 Eye-Tracking Gaze

In this section we present our results based on the output eye-gaze data provided by the eye tracker: time in areas of interest and eye-gaze pattern.

7.2.1 Time in Areas of Interest

One question that we aim to respond to in this section is: how are the density and sparsity of familiarity and trust networks correlated with the gaze-patterns of our subjects?

Firstly, we recall that members in Group2 are more connected than members in Group1: familiarity and trust networks are denser as it was inferred in Figure 3. In Figure 13 we see that for all interfaces and any number of members, users in Group2 looked less at AOI Songs and spent more time by looking at AOI Users and AOI Ratings: 38% vs. 28% for all interfaces, 24% vs. 15% for all members, 42% vs. 25% for closest 5 members and 53% vs. 47% for closest 2 members. This implies that they were mainly interested in seeing and analyzing other members' ratings. Members in the first group behaved more individualistic, by looking most at AOI Song and spending less time in the other 2 AOIs. This observation can be summarized as follows: the more connected a group is through familiarity and trust, the more interested its members are to discover other members' music preferences.

The results also show that, within group, participants looked less at AOI Songs for the first 8 interfaces, i.e. when all other members and their ratings where displayed. This is an expected behavior since the most information was included in the first part of the experiment and users needed more time to scan through the names and ratings: the area of the 2 AOIS (AOIU and AOIR) is the biggest.

Another important finding is that for the interfaces containing 2 closest members, participants in both groups looked approximately half of the time at the top (AOI Songs: grey rectangle in Figure 9) and half at the bottom of the interface (AOI Users and AOI Ratings): 53% for Group1 and 47% for Group2 for AOI Songs.

We compare the time in AOI Users with the time in AOI Ratings within group. Over all interfaces time in AOI Users is longer than time in AOI Ratings as follows: 34% vs. 28% overall, 39% vs. 37% for all members, 32% vs. 26% for closest 5 members and 29% vs. 18% for closest 2 members.

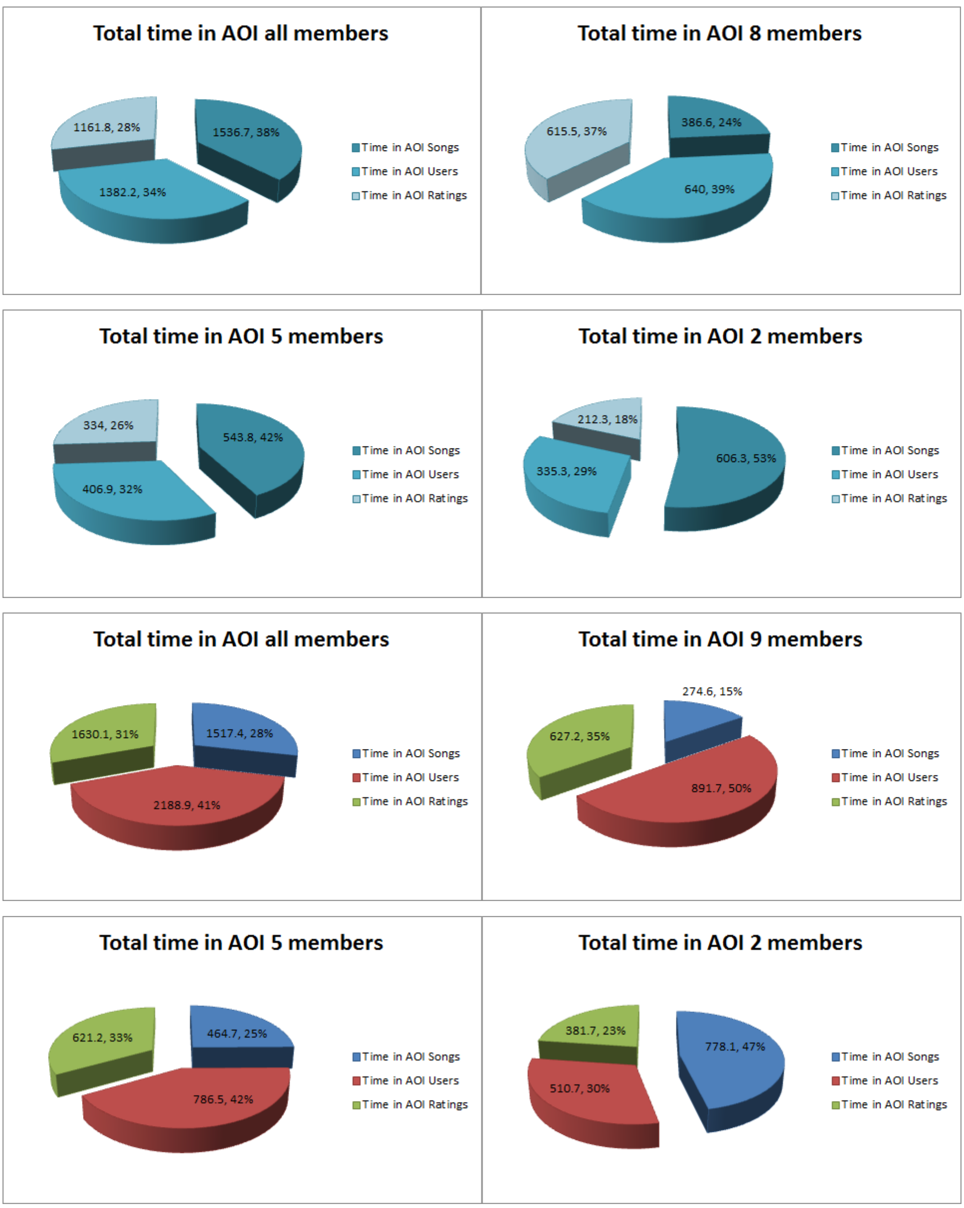


Figure 13. Time in AOI distribution: Group1 (top 4 charts) and Group2 (bottom 4 charts)

The relative percentage difference between AOI Users and AOI Ratings is less for all 8 members in Group1 and for 2 closest members in Group2. Turning back to the network structure we find this result very encouraging. We explain it by the fact that the more distant the relationship between users is (Group1) the more they are inclined to look at the whole list of members rather closest ones.

7.2.2 Eye-Gaze Pattern

The eye-gaze pattern of each participant offers a great detail of information highlighting the attention a subject paid to a certain part of the interface before making a decision.

As it can be inferred from the below figures there are two types of correspondences which help understand the importance of information displayed to each participant through an objective measurement: horizontal lines – between users (left) and their ratings (right) – and vertical lines – within AOI Users or AOI Ratings. The (almost) horizontal correspondences are the ones which are important for our analysis as they contain the relevant information in the form of the association that a subject made (Figure 14). By analyzing a vertical line we record both the other member's name (to the left) and his rating (to the right).

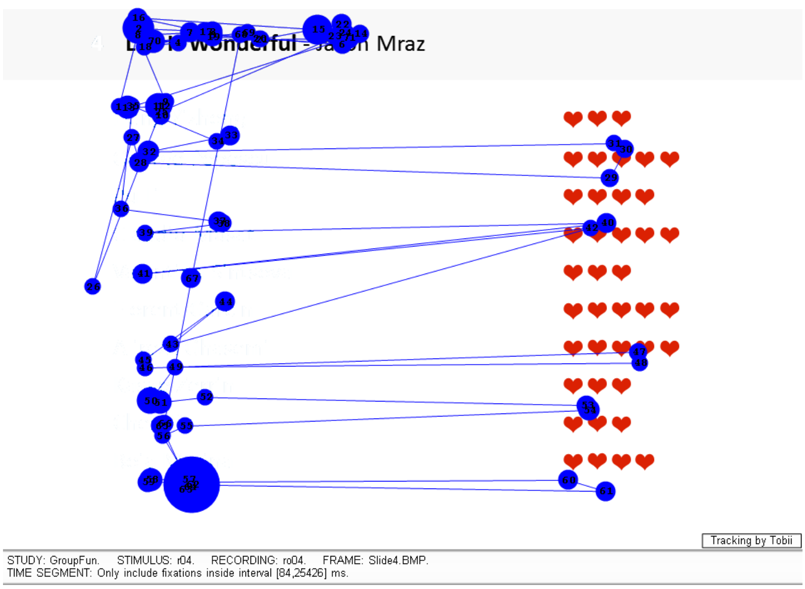


Figure 14. Example of eye-gaze pattern

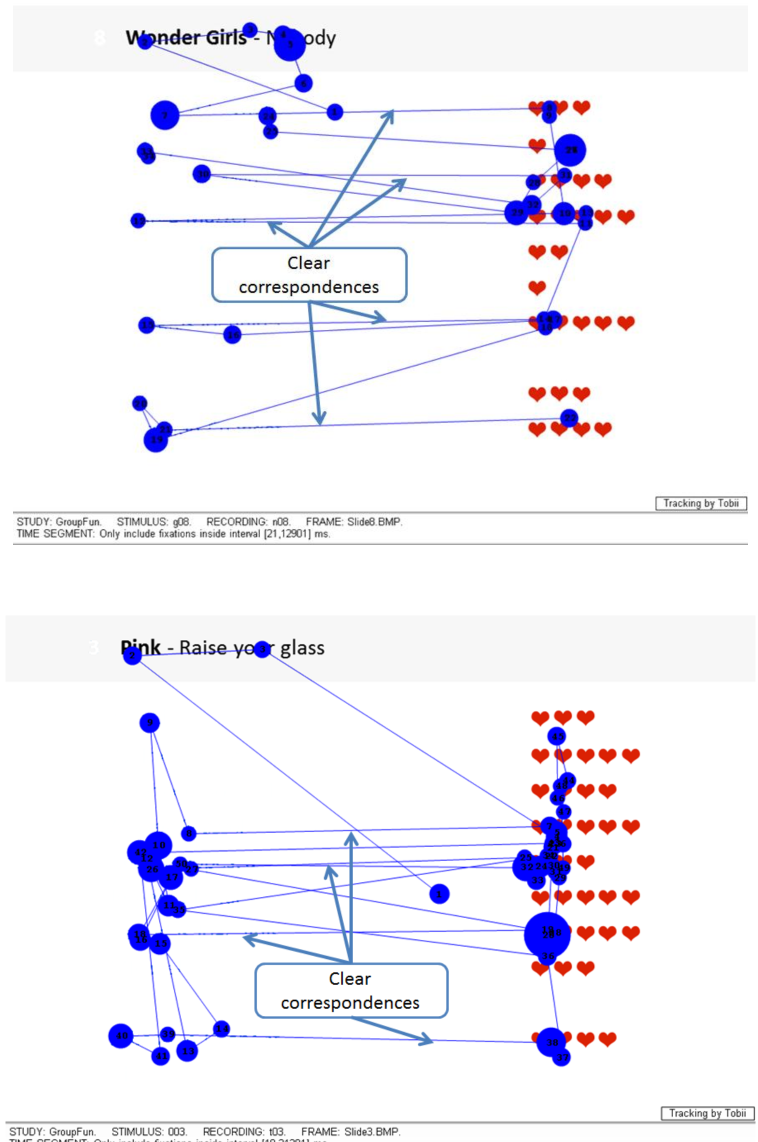


Figure 15. Examples of clear horizontal correspondences.

In Figure 15 are marked clear horizontal correspondences between users and their ratings. Out of the entire number of subjects to the left (as mentioned before, hidden for privacy concerns) the current participant in Figure 15 (top) – member of Group 2 (there are 9 users and their ratings displayed) - looked at only 5 other members' ratings and drew associations with their ratings in evaluating his own rating. We consider only the correspondences which connect names and ratings on the same raw (we call them "clear correspondences"), such as the ones in the figure above. Clear correspondences can vary with a small angle deviating from the horizontal line. The screen-shot in Figure 15 (bottom) shows the eye-gaze of another participant. This subject looked more to the bottom of the interface and created 4 horizontal correspondences that were recorded. In general, not only that we obtain very good eye-gaze recording and fixation times but also most of the outputs are similar to those from Figures 14 and 15.

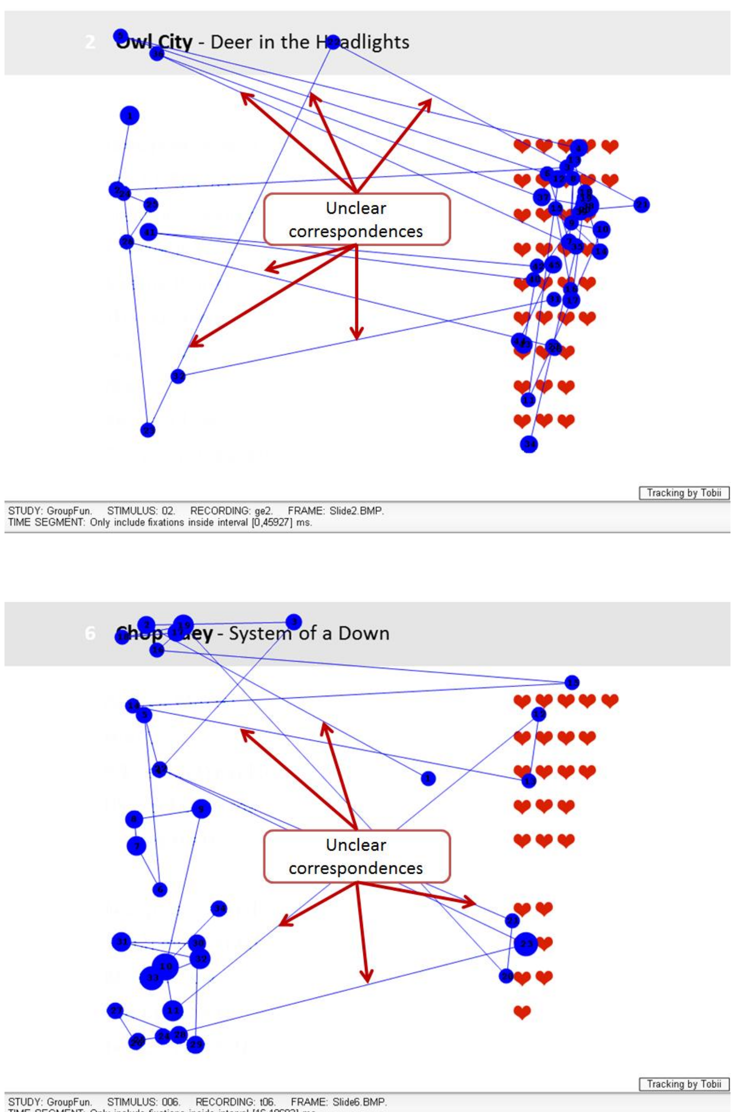


Figure 16. Examples of unclear visual correspondences

However, not all eye gaze patterns were very clear. We noticed some outputs in which users looked:

(1) mostly (90% of the time or more) at the song's name;
(2) at parts of the interface displaying no information;
(3) from top to bottom on both users and ratings columns and draw few associations;
(4) in a non-uniform (even chaotic!) manner, skimming from one corner of the interface to the next.

In this case fixation times do not correspond well to the names on the left and corresponding lines are non-horizontal but skewed. Consequently, we state that the participant combined the 3 AOIs drawing associations from AOI Song to AOI Users (top names) and from AOI Song to AOI Ratings (middle and bottom hearts) which is undesired for a clear evaluation. Here we cannot conclude anything about group influence for any of the members on the left but can only use the fixation times for each of the 3 AOIs.

In Figure 16 the correspondences include skewed lines between different information: (1) song and ratings and (2) users and ratings. In general participants skipped rating information more than name information, as it can be inferred from the computed time in AOIs. Vertical lines are denser in the left side (AOIU) than in the right side (AOIR) but the focus corresponds to the horizontal lines which are clearly marked.

Eye-gaze patterns of our participants contained a mixture of clear and unclear correspondences. Such associations are presented in Figure 17. The skewed lines follow an oblique trajectory whereas the clear correspondences follow an almost horizontal path. The eye-gaze pattern includes the complexity of eye-movement on the screen showing both associations and deviations.

We mark the clear horizontal correspondences per participant by associating them with the other group members. Such saccades include the participants' mental model of decision making by filtering the relevant information from the interface and considering others' ratings in own valuation.

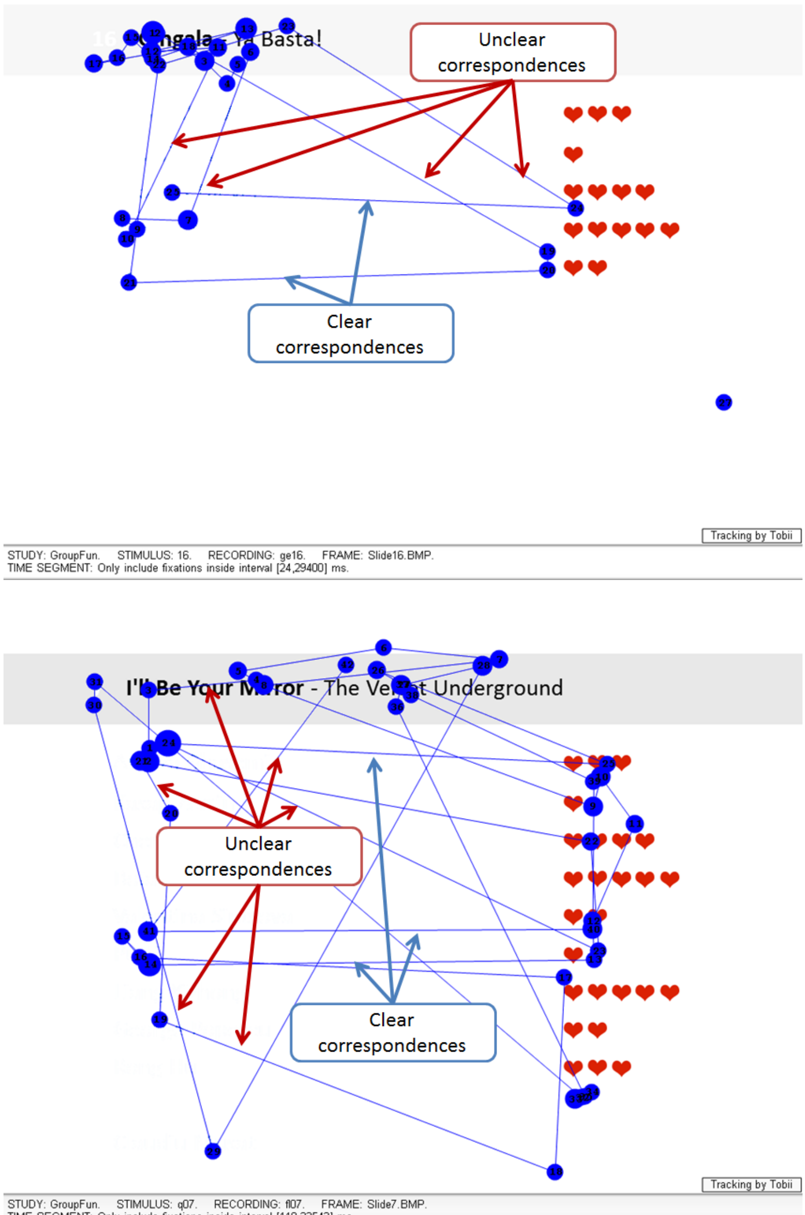


Figure 17. Examples of complex eye-gaze patterns

7.3 Preference Correlation

We used the Pearson correlation score reflecting the degree to which 2 variables are related to determine subjects who are closest preference-wise. In equation (1) r represents the rating submitted for each of the 24 songs, r_i is the rating submitted by participant i and $\bar{r}$ is the average of ratings. Largest correlation values are marked in bold in Table X for each participant. We mapped this with the familiarity and trust network to observe if the subjective evaluation of familiarity and trust corresponds with objective ratings.

$$\rho_{i,j} = \frac{\sum_{i=1..8/9} \left(r_i - \bar{r}\right)\left(r_j - \bar{r}\right)}{\sqrt{\sum_{i=1..8/9} \left(r_i - \bar{r}\right)^2} \sqrt{\sum_{j=1..8/9} \left(r_j - \bar{r}\right)^2}} \quad (1)$$

As an example, in the first group, P1 thinks is closest to P4 (highest familiarity score) but the ratings show him closest to P8 (preference correlation values). Reciprocally, P4 thinks to be closest to P1 (again, highest familiarity score) but rating correlations prove her to have closest tastes to P9. Another case shows P7 to know most P5 and P9's music tastes. Once again the correlation coefficient shows that he likes music similarly to P8.

In the second group P19 is the "standard" rating reference for all other members. P11, P13, P14, P16, P17 and P18 all have highest rating correlation with P19. Furthermore, P18 and P19 stated to have known each-others' preferences the most. Another interesting result is that P12 knows P15's tastes and these tastes are also highly correlated.

Correlation between music preferences has been recently used in recommendation systems to produce more accurate suggestions inferring people's tastes from similar profiles. The aggregation method is named collaborative filtering (CF). This method makes automatic predictions about the interests of a user by collecting the preferences or tastes information from many users. The underlying assumption of CF is that if a user A has the same preference as user B for (a) certain item(s) then A is likely to "inherit" B's preference for a different item. In many cases, such as music, movies, books, travelling, collaborative filtering proved to be a better technique than the simpler approach of averaging scores based on votes or ratings.

At the end of the experiment our participants were curious to find out which other members have similar tastes with them. They expressed their interest to follow these preferences based on similar ratings and eye-tracking correspondences. We took into account the first factor (the 24 ratings per group) to compute the Pearson correlation values for all 24 songs between each pair of users in the two groups. The higher the value the more alike individuals are in terms of sharing similar music preferences. We made this information available to all participants at the end of the study.

In the two tables below (Table X), P1 shares most similar music preference with P2 (0.73 correlation value) and most different preferences with P8 (-0.6). Similarly, P5 is closest to P3 (0.4) but has opposite music tastes than most other members (highest negative values are P1: – 0.49 and P2: -0.67). In the second group P13 has most similar music preferences with P11 (0.47), P14 (0.43) and P18 (0.41).

Table X. Pearson correlation values based on the 24 eye-tracking songs' ratings: Group1 (top) and Group2 (bottom)

	P2	P3	P4	P5	P6	P7	P8	P9
P1	**0.73**	-0.54	0.22	0.18	-0.49	0.20	-0.60	-0.15
P2		-0.43	0.32	0.09	-0.67	0.17	-0.48	-0.10
P3			-0.06	0.22	0.40	0.30	**0.70**	0.29
P4				-0.02	-0.19	0.47	-0.16	-0.31
P5					-0.19	-0.04	-0.16	-0.05
P6						-0.02	**0.61**	0.40
P7							0.12	0.11
P8								0.18

	P11	P12	P13	P14	P15	P16	P17	P18	P19
P10	0.21	0.38	0.25	0.19	-0.18	-0.20	0.33	0.47	0.19
P11		0.02	0.39	0.47	0.17	0.14	0.16	0.38	0.34
P12			-0.01	0.13	**0.59**	-0.32	0.14	0.42	-0.30
P13				0.43	-0.03	-0.05	0.10	0.41	0.38
P14					0.19	0.10	-0.23	0.39	-0.10
P15						-0.24	-0.18	0.40	-0.40
P16							0.02	-0.27	-0.10
P17								0.02	0.25
P18									0.00

We plot the two tables into two un-oriented graphs (Figure 18) using Gephi and compare them with the familiarity and trust networks from Figure 3. In the center of two graphs below are the participants with highest correlation with others, i.e. the ones closest to the groups "average" or common preference.

In Group1 participants P1 and P3 are in the center whereas in Group2 participants P11, P18 and P19 obtained highest correlation values with other members. This cross-correlation shows the objective measurements of the individual preference elicitation vs. the first graphs which denote the subjective measurement values of familiarity and trust. Thus, P1 and P3 positioned in the center of the green graph and P11, P18 and P19 positioned in the center of the red graph are closest to the group decision even though they have not necessarily received highest familiarity or trust scores.

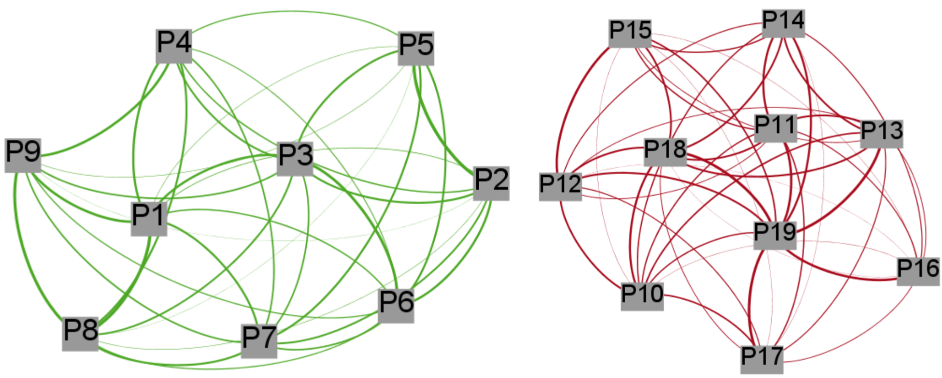


Figure 18. Pearson correlation networks for the two groups

Comparing the familiarity and trust networks in Group1 with the Pearson correlation (re-scaled from [-1, 1] to [0, 5]) music-preference network we observe that the overlay between values is quite small. P4, for instance, mentioned to know P1's music preference but he has most similar preferences as P7 whereas all other correlation values are close to 0. He also trusted P1 the most and P2, P3, P5, P6, P8, P7 P8 and P9 equally. If we take the example of P12 from the second group, he is most similar music-wise with P11, P15 and P18 (Table X). However, he trusts most P10, P15, P17 and P18 and knows most P13, P15 and P18's preferences. We can conclude that only to a small extent the music-preference correlation network match the familiarity and trust networks.

7.4 Ratings Change

In this section we measure the change in ratings for the two conditions: individual rating in GroupFun and group rating in the eye-tracking studies. Based on the changes we compute a social influence score per group. We further elaborate our analysis after excluding own contributed songs and derive symmetry and transitivity relationships. In addition to our experiment setup we compute the change rate for the consensus and divergence conditions.

7.4.1 Ratings in Individual and Group Interfaces

Not only that the social interfaces that we designed helped us record fixation times and understand the meaningful information per participant but they also proved to support the same satisfaction for both groups (Table XI). The difference from Table III before is that here we consider the same 24 songs which were rated and displayed on the interfaces instead of all the contributed group songs: 39 for the first group and 47 for the second one. The group satisfaction score is again the average of all users' ratings.

The "Average Before" condition in the third column of the table below includes the average values of all ratings that a group submitted before seeing the eye-tracking interface. The "Average Interface" column includes the ratings displayed in the interfaces. The only difference between the two groups in this case is for the 8/9 members condition. Since one interface had one extra member and his rating we computed the closest average values with about the same standard deviation: average of 3.43 for the smaller group (1.14 standard deviation) and average of 3.50 (1.13 standard deviation) for the larger one. Finally, the last column "Average after" includes the average rating values per group after the participants saw the group interfaces.

Comparing columns 3 and 5, it yields that generally both groups slightly reduced their ratings showing less satisfaction. For all interfaces the average rating in Group1 dropped from 3.21 (SD = 0.64) to 3.15 (SD = 0.41) while in Group2 it also dropped from 3.53 (SD = 0.57) to 3.44 (SD = 0.31). By opposition, a significant increase can be observed for the last condition: "2 closest members". Here both groups increased their average ratings from 3.10 to 3.27 (Group1) and from 2.93 to 3.33 (Group2). This means that subjects were more open and gave higher ratings when seeing the ratings from most trusted peers.

Table XI. Satisfaction comparison per group, sets of interfaces and experiment phase

Interface	Group	Average Before	Average Interface	Average After
All	**1**	3.21 (SD = 0.64)	3.52 (1.24)	3.15 (SD = 0.41)
All	**2**	3.53 (SD = 0.57)	3.55 (1.24)	3.44 (SD = 0.31)
8	**1**	3.11 (SD = 0.79)	3.43 (1.14)	3.05 (SD = 0.38)
9	**2**	3.69 (SD = 0.57)	3.50 (1.13)	3.46 (SD = 0.50)
5	**1**	3.05 (SD = 1.02)	3.65 (1.17)	3.11 (SD = 0.63)
5	**2**	3.75 (SD = 0.84)	3.65 (1.17)	3.52 (SD = 0.50)
2	**1**	3.10 (SD = 1.01)	3.50 (1.41)	3.27 (SD = 0.49)
2	**2**	2.93 (SD = 0.72)	3.50 (1.41)	3.33 (SD = 0.55)

Despite the high standard deviation of ratings displayed in the interfaces (all bigger than 1), we report higher preference alignment after seeing the group interfaces for both groups. The numbers in the brackets show very low standard deviation values (max. of 0.63) for all conditions and both groups. Comparing this with the standard deviation from the 3rd column in which values are much higher we find that, despite the fact that individual preference diverged more when taken individually, subjects' ratings converged more when shown the same reference, i.e. the eye-tracking interfaces.

One final observation targets the average ratings in the interface. As it can be observed from the "Average Interface" column the average rating in all the 24 interfaces was very balanced: 3.43 smallest average for "8 other members" in Group1 and 3.65 largest average for both groups for "5 closest members". With this design we aimed at not influencing our subject's decisions through ratings only, but through social pressure coming from most trusted members.

7.4.2 Change Rate

We computed a normalized group influence index per participant measuring the difference in ratings which each subject submitted while being exposed to the individual interface vs. the group interface.

For each group member we compute an average rating for the GroupFun experiment in Phase I as a measure of individual satisfaction (Equation 2). We normalize this by the number of rated songs since this number is different among users. U is the set of all users (participants or subjects), R is the set of all ratings and S the set of all songs.

$$satisfaction_i = \frac{\sum_{1 \le j \le S} r_{ij}}{\# r_i}, \forall i \in U, r_{ij} \in R, |S| = 39 / |S| = 47 \tag{2}$$

Similarly, we compute a satisfaction score for the eye-tracking experiment (starred) considering only the ratings given by the 24 songs in the interface (equation (3)):

$$satisfaction_i^* = \frac{\sum_{1 \le j \le 24} r_{ij}^*}{\# r_i^*}, \forall i \in U, r_{ij}^* \in R^* \tag{3}$$

$\# r_i$ is the number of rated songs for each participant out of the 39 and 37 songs, respectively. For the second equation $\# r_{ij}^* = 24$ as all participants rated all 24 songs.

Next, we compute a normalized group influence index per participant measuring the difference in ratings he submitted while being exposed to the individual interface compared with the group interface, only for the same songs (that were rated twice).

$$\Delta_i = \sum_{1..j} \left| r_{ij} - r_{ij}^* \right|, \forall i \in U, r_{ij} \in R, r_{ij}^* \in R^* \tag{4}$$

$$\varepsilon_i = \sum_{1..j} r_{ij} - r_{ij}^*, \forall i \in U, r_{ij} \in R, r_{ij}^* \in R^* \tag{5}$$

In Table XII we denote the following:

- Avg_i: average rating for the individual, GroupFun interface – 24 songs;
- Avg_g: average rating for the group, eye-tracking interface – 24 songs;
- t: number of songs rated twice (in both interfaces) in 3 categories: “+”, “-“ and “0”;
- +: number of positive changes – increase of ratings;
- -: number of negative changes – decrease of ratings;
- 0: number of neutral changes – ratings kept the same;
- Δ: absolute rating difference;
- ε: signed rating difference;
- c: change rate percentage.

In the table below we compare the ratings and change rate only for the 24 songs displayed for the second part of the experiment. One first observation is that participants did indeed change their ratings at a very high rate (40.15% in Group1 and 50.19% in Group2). Once again, the more connected group, Group2, changed ratings most. We marked in bold the participants who had a high relative change rate. One outlier is participant P2 who did not change any of his ratings for the 7 re-rated songs. Now, if we sum the values from the two "change" columns, which are the "+" and "-" columns, it yields an average of 45.17% of changed ratings out of the total number of re-ratings. Participants in Group1 rerated 152 songs out of which 42 where an increase in values, 25 a decrease and 85 where kept the same. In Group the 10 participants submitted 271 new ratings out of which 63 where increases, 64 decreases and 144 where kept the same.

Table XII. Rating change statistics

	Avg_i	Avg_g	t	+	-	0	Δ	ε	c
P1	3.29	**3.58**	17	8	6	3	21	7	**82.35%**
P2	**4.00**	3.33	7	0	0	7	0	0	0.00%
P3	**3.63**	3.42	16	5	2	9	9	3	43.75%
P4	3.08	**3.29**	24	10	7	7	29	5	**70.83%**
P5	**3.78**	3.63	23	0	2	21	2	-2	8.69%
P6	**3.50**	2.88	8	1	1	6	2	0	25.00%
P7	2.14	**2.54**	22	9	3	10	16	10	**54.54%**
P8	2.24	**2.54**	21	5	2	14	7	3	**33.33%**
P9	**3.21**	3.17	14	4	2	8	7	3	42.86%
Group1	**3.21**	**3.15**	**152**	**42**	**25**	**85**	**93**	**29**	**40.15%**
P10	3.63	**4.14**	7	2	1	4	6	0	42.85%
P11	3.54	**3.56**	9	1	3	5	4	-2	44.44%
P12	3.96	**4.21**	14	1	4	9	5	-3	35.71%
P13	3.83	**4.27**	11	5	0	6	8	8	45.45%
P14	3.5	**3.57**	7	1	2	4	4	-2	42.86%
P15	2.96	**3.56**	9	3	2	4	6	2	**55.56%**
P16	3.13	**3.18**	11	2	6	3	11	-5	**72.73%**
P17	**3.29**	2.54	13	2	7	4	20	-14	**69.23%**
P18	3.25	**3.41**	17	2	3	12	5	-1	29.41%
P19	**3.33**	2.91	22	3	11	8	16	-8	**63.64%**
Group2	**3.44**	**3.54**	**120**	**22**	**39**	**59**	**85**	**-25**	**50.19%**
Total	**3.32**	**3.34**	**271**	**63**	**64**	**144**	**178**	**4**	**45.17%**

The signed difference and absolute difference between the two arrays of ratings, r_{ij} and r_{ij}^{*}, are computed in order to determine the number of changes: positive: “+” (increase), negative: “-” (decrease) and neutral: “0” (same).

In the second column of Table XI we marked in bold the higher average satisfaction of the two conditions per participant: individual (second column) and group (third column). We report that only 7 (out of 19) participants were more satisfied with the music they rated in the first condition. All others increased their ratings. Overall, the group satisfaction varies only slightly: 3.15 for the group interface for Group1 compared with 3.21 and 3.54 compared with 3.44.

Additionally, both Δ and ε help understand the variations of ratings. The minus sign correspond to a decrease of ratings compared with the individual interface used as baseline. The decrease and increase may correspond to a large or small variation in rating change according to each rating update. Despite the fact that the number of re-ratings vary significantly per participant: minimum of 7 songs in both groups and maximum of 24 in Group1 (P4) and 22 in Group2 (P19) we report a large change rate variation across participants.

Another important result is that members in the first group have achieved about the same satisfaction in both conditions - 4 members were more satisfied and 5 less satisfied - whereas in the second group 8 out of 10 members have improved their ratings overall - only P17 and P19 were less satisfied than before.

Across all participants we measure an average group change rate of 45% - average of all 19 c-values. Overall, the more songs people rate the higher the group influence and social alignment - the bold percentages from the right column mark group influence higher than 50%. P1 (82.35%), P4 (70.83%) from Group1 and P16 (72.73%), P17 (69.23%), P19 (63.63%) and P15 (55.56%) are the participants that changed the most. This phenomenon is related to one’s memory. If one person re-rates a few songs in a short time period then he would be more inclined to re-state the same evaluation rating again.

7.4.3 Group Influence Score

Next, we consider the following 2 functions:

$$\bar{r}_i = \frac{\Delta_i}{\# r_i}, \forall i \in U, r_i \in R \tag{6}$$

$$\eta_i = \frac{\sum_{1..j} \left(\# r_{ij}^{+} + \# r_{ij}^{-} \right)}{\sum_{1..j} \left(r_{ij}^{+} + r_{ij}^{-} + r_{ij}^{0} \right)}, \forall i \in U, j(i) \in R \tag{7}$$

The first one computes the average rating difference (on a 1 to 5 scale) for the absolute rating difference whereas the second one is a value between 0 and 1 measuring the number of rating changes vs. the total number of ratings. The variable from Equation 7 represents the normalized group influence score.

For the computation of the group influence score we rely on the following data: the same songs re-rated per participant and number of changes in ratings. The results from Figure 19 show the difference for the group interfaces compared with the ratings recorded in GroupFun. This chart allows us to recognize, on one hand, participants who changed their ratings the most: P1, P2 in the first group and P17, P16 and P10 in the second (higher than a score of 0.8 of 5 on average, per song) and, on the second hand, participants who did not change almost at all their preferences: P2, P5, P6 and P8 in Group1 and P18, P12 in Group2 (lower than a score of 0.4 out of 5 on average). Noticeably, participant 2 didn't change any of his ratings for the 7 songs he previously rated! To understand this change we correlate the group influence index with eye-tracking data in the following section.

As it can be inferred from Figure 20 the change rate (Table XII) for both groups is very high suggesting that individuals were strongly influenced by their peers' ratings. An average η for Group1 is 0.4 and for Group2 0.5 meaning that, on average, each participant in Group1 changed his previous rating with 0.54 (on a 1 to 5 scale) and each participant in Group2 changed each of his ratings with 0.72. Indeed social influence in the second group is more homogenous given the groups higher connectivity compared with the first group (Figure 3). In Group1 some participants tended to have a very strong preference which they care about and would not adapt to the group decision. In the next section we compare this explicit group influence with the objective eye-tracking correspondences measuring the other members which participants paid most attention to. Data shows that individuals do not have fixed preference values but adapt to the group.

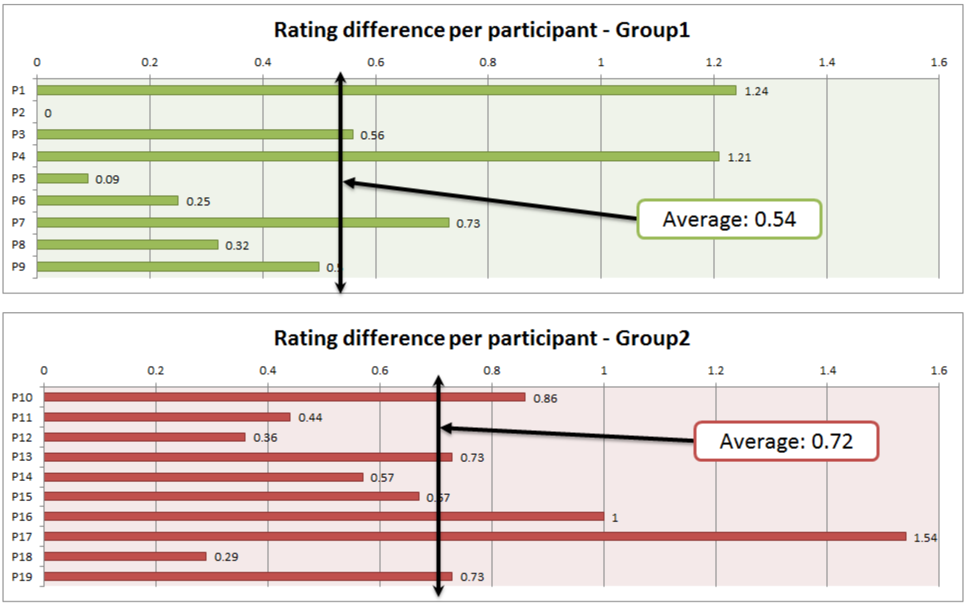


Figure 19. Rating difference per participant in both groups

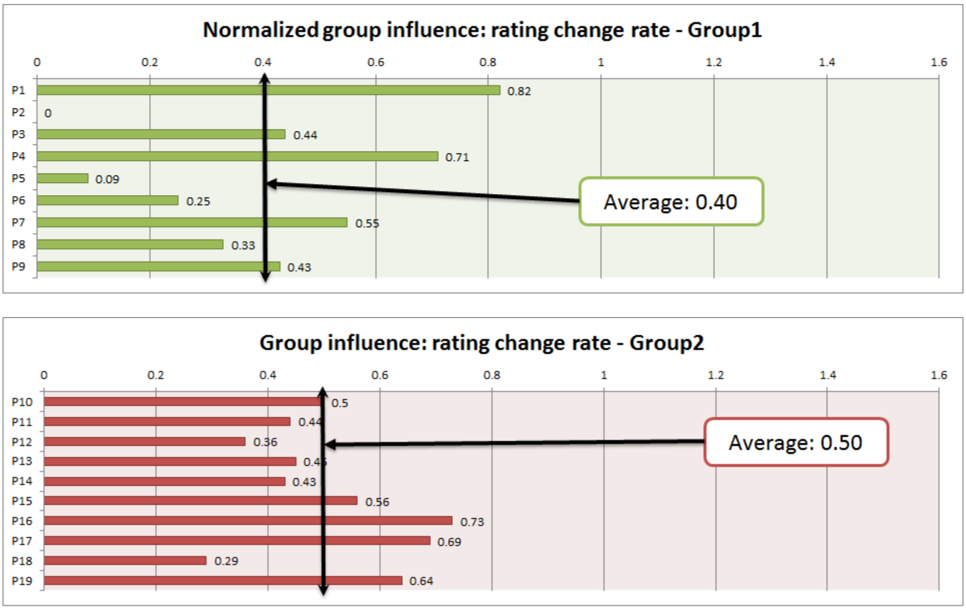


Figure 20. Group influence score for both groups

7.4.4 *Quantifying Change Rate*

Another dimension we explore understanding rating changes is categorizing rating differences into "1", "2", "3" and "4" (maximum) point(s) corresponding to the absolute difference between the initial submitted ratings (in GroupFun) and the updated ones (during the eye-tracking experiment). "1" represents small rating differences such as "3-2" or "4-5" whereas 4 corresponds to the maximum difference which is either "1-5" or "5-1". In Table XIII each column corresponds to the number of differences between the initial and the updated ratings.

Table XIII. Quantified rating changes: 1, 2 and 3-points differences

	1-point	2-points	3-points	Out of max. = K
P1	9	3	2	17
P2	0	0	0	7
P3	5	2	0	16
P4	8	6	3	24
P5	2	0	0	23
P6	2	0	0	8
P7	8	4	0	22
P8	7	0	0	21
P9	5	1	0	14
Group1	**46**	**16**	**5**	**67/152**
P10	1	1	1	7
P11	4	0	0	9
P12	5	0	0	14
P13	5	0	0	11
P14	2	3	0	7
P15	4	1	0	9
P16	6	1	1	11
P17	3	3	1	13
P18	5	0	0	17
P19	12	2	0	22
Group2	**44**	**12**	**3**	**61/120**
Total	**90**	**28**	**8**	**128/272**

This table represents a greater detail of the change rate percentages calculated in Table XIII. Using these divisions we understand better where the change comes from comparing the variation in ratings and the total number of ratings for all re-rated songs (K).

For each participant in each group we count the number of times each rating difference falls into the “1”, “2”, “3” and “4” categories corresponding to “small”, “medium”, “large” and “very large” changes. In total, for both groups, we obtained 128 changes out of which 90 were 1-point (70% “small changes”), 28 were 2-point (22% “medium changes”), 8 were 3-point (6% “large changes”) and 2 were 4-point rating changes (2% “very large changes”). Indeed, in 2 cases the group influence was the highest and made one individual (both completely opposite changes came from the same subject) change his rating from “1” (“strongly dislike the song”) to “5” (“strongly like the song”) and vice-versa. This is another important result for measuring group influence.

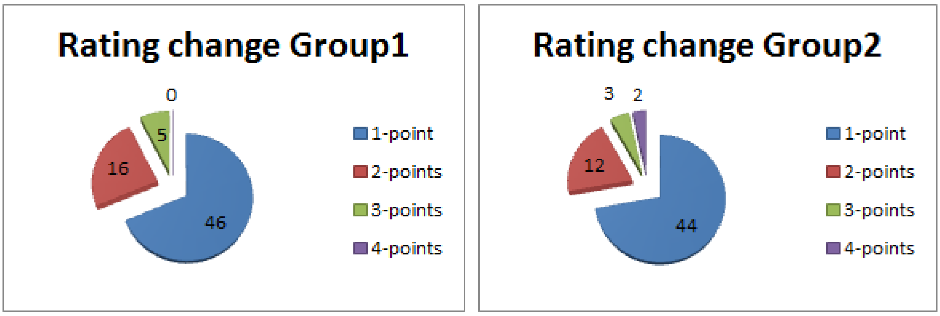


Figure 21. Rating change per group for 1, 2, 3 and 4-point(s) differences

The two groups exerted similar change behavior. In total we noted 67 rating changes out of 152 songs rated twice in Group1 (44% change rate) and 61 rating changes out of 120 songs rated twice in Group2 (51% rating change). The rating differences between groups also show similar results: 46 vs. 44 “1”-point rating changes, 16 vs. 12 “2”-point changes, 5 vs. 3 “3”-point changes and 0 vs. 2 “4”-point changes in Group1 and Group2, respectively. The numbers in Tables X and XI are slightly different because 2, 3 and 4-points change rate are not quantified the same as 1-point changes with respect to change rate. We expected to obtain higher numbers for the second group given the overall social alignment and the results obtained from the eye-tracking data. Members’ preferences prove that this is not the case and that variations are small. A more detailed analysis considering each user separately yield that individuals exert very diverse preference behavior. The vast majority of subjects changed their rating only slightly which corresponds to a more general psychological result that people have intrinsic preferences which they value. They are influenced by their peers and the context but are considering changing their preferences only to a small extent.

To summarize this section, our data proves that individuals do not have fixed preference values that they conform to but they adapt to group suggestions.

7.4.5 Change Rate With and Without Own Songs

Given that individuals tend to overstate their preferences for their own contributions we decided to remove from our analysis self-uploaded songs. For both groups the difference is presented in Figure 22 below and includes only 6 “1”-point rating differences for self-uploaded songs; the other differences are the same

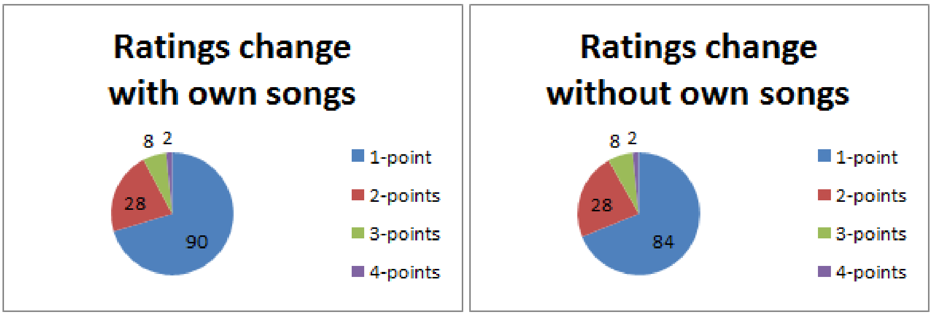


Figure 22. Rating change for both groups

7.4.6 Social Influence and Group Size

The following analysis is based on the selection of all ratings given to songs which do not include self-contributions. For each type of eye-tracking interfaces displaying: (1) all other group members, (2) 5 closest members, and (3) 2 closest members we compute the average change rate per group as well as for all participants. The aim is to understand if individuals align more when they see many opinions or when they see fewer ones from people they trust most. Our hypothesis was that they would change more when they face the preferences of 2 closest members and re-state the same preferences when facing everyone else in the group.

The table below (Table XIV) summarizes the change rate for the 3 conditions stated before per group. The last row is the weighted average between Group1 and Group2 taking into account the group size: 9 members for Group1 and 10 members for Group10.

Table XIV. Change rate per group for the 3 types of interfaces

	All other members	5 closest members	2 closest members
Group1	**46.20%**	35.93%	38.02%
Group2	44.77%	**58.71%**	**56.96%**
Both groups	45.37%	49.12%	48.99%

First of all, when comparing the two groups among themselves we notice that the members of the second group adapted more to the group decision as the number of members in the interface was reduced. This validates our hypothesis: Closest 5 and Closest 2 have similar change rate scores: 58.71% and 56.96%, respectively. However, in the first group the change rate decreased from the first group of 8 interfaces to the last ones. This proves that members changed their preferences less when seeing fewer opinions, even if those opinions came from the members that they trusted the most!

This is opposite to the trend noticed for Group2, the more connected group. This group changed least when seeing all other members' preferences: 44.77% and most when seeing 5 closest members (58.71%) and 2 closest members (56.96%). The main reason for this phenomenon we associate with the familiarity and trust networks. Group1 is the one in which people are more isolated and adapt less. Consequently, this result is in agreement with this group's characteristics. The situation in Group2 is reversed: in this connected group individuals adapt more, gaze more at others' preferences and only slightly decrease their overall satisfaction. The change rate further proves that they also adapt more when they see their 5 and 2 closest members' preferences. It is also interesting to note that the "starting point" of both groups is about the same: 46.20% vs. 44.77% (Table XIV) change rate for all other group members. Indeed, social relationships are beneficial for the individual decision making process in the second group. A bar-chart presenting the results from Table XIV is included in Figure 23 below.

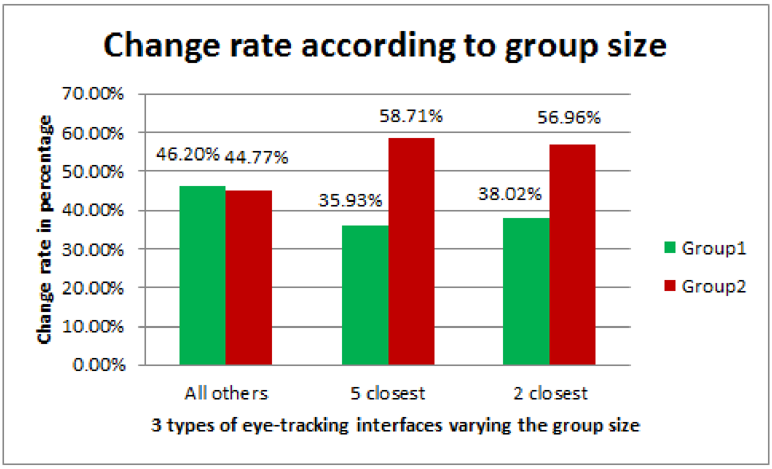


Figure 23. Change rate and group size comparison per group

7.4.7 Change Rate for Consensus and Divergence

In the following we analyze the change rate for consensus vs. divergence conditions in both groups. The data from the Table XV below show that the second group, in general, changed more for all the (sub-) conditions. Furthermore, individuals changed most when seeing 2 members' preferences, then 5 others and finally all others in the strongly connected group (Group2) under both the consensus and divergence conditions. This detailed result was obtained after investigating the change rate percentages from Table XIV. For our analysis it is encouraging to obtain an increase in the change rate as the group size decreases.

In Group2, for consensus we obtained: 63.63% change rate for all other members, 65.00% change rate for closest 5 members and 70.00% for 2 closest members. We report even greater relative differences for the divergence condition: 45.83% change rate for all other members, 55.85% for 5 closest members and 68.33% change rate for 2 closest members.

In Group1 the situation is more complex. First of all, under both the consensus and the divergence condition, overall participants changed the same. For the consensus condition subjects changed most when seeing 5 closest members (56.48%), then all other members (44.79%) and finally 2 closest members (only 33.33%!). However, when facing divergent ratings they changed most when seeing 2 closest members (66.66%), then all other members (60.18%) and least for 5 closest members (52.77%).

Table XV. Change rate comparison for consensus and divergence in the two groups

Change Rate – Group1: 44.04%		
Group Size / Variable	**Consensus**	**Divergence**
All other members	2 – 44.79%	2 – 60.18%
5 closest members	1 – 56.48%	3 – 52.77%
2 closest members	3 – 33.33%	1 – 66.66%
Overall Consensus / Divergence	41.59%	46.61%
Change Rate – Group2: 52.92%		
Group Size / Variable	**Consensus**	**Divergence**
All other members	3 – 63.63%	3 – 45.83%
5 closest members	2 – 65.00%	2 – 55.83%
2 closest members	1 – 70.00%	1 – 68.33%
Overall Consensus / Divergence	49.71%	54.07%
Overall: 48.71%		

We expected individuals to change less when consensus was presented to them and more when the variance of others' preferences was large. However, the data shows that this is the case only for the first group. Despite close scores, the change rate in both groups shows that individuals changed their preferences more when they saw divergence.

7.5 Absolute Rating Difference

We noticed that the rating differences among participants account for the social pressure of other group members. Thus, we investigate in more detail the rating differences. In this section we measure the rating difference between initial (individual) and eye-tracking (group) conditions for the same set of 24 songs. First of all, we select both vectors of ratings and compute the unsigned (absolute) difference between them (Equation 8). Then, we sum the difference for each song per participant and divide by the total number of songs rated twice in the two conditions (Equation 9). In this way we obtain the D-values for all participants that we use to compute a group D-value for Group1 (Equation 10) and Group2 (Equation 11) and compare D-values among experiment conditions.

$$d_{i,j} = | r_{i,j} - r^*_{i,j} |, \forall i \in U, j \in S, r_i \in R, r_i^* \in R^* \quad (8)$$

$$d_i = \frac{\sum_j d_{i,j}}{k_i}, \forall i \in U, j \in S, k_i \in R \cap R^* \quad (9)$$

$$\overline{d}_i^1 = \frac{\sum_{i=1}^{9} d_i}{9}, \forall i \in U, r_i \in R, r_i^* \in R^* \quad (10)$$

$$\overline{d}_i^2 = \frac{\sum_{i=10}^{19} d_i}{10}, \forall i \in U, r_i \in R, r_i^* \in R^* \quad (11)$$

We populate the table below with all the D-values computed for all sets of conditions. The generic D-value corresponds to the value from Equation 12 below, which calculates the average of all D-values across all 19 participants.

$$\overline{d} = \frac{\sum_{i=1}^{19} d_i}{19}, \forall i \in U \quad (12)$$

Table XVI. Absolute difference values for opposing experiment conditions

Overall	Trust vs. Popularity	Consensus vs. Divergence	Fixed Divergence	Trust vs. Popularity
All members	0.586	0.586	**All members**	0.623
5 members	0.652	0.659	**5 members**	0.565
2 members	0.738	0.729	**2 members**	0.756
All interfaces	**0.654**	**0.654**	**All interfaces**	**0.646**

Fixed Familiarity	Cons. vs. Div.	Fixed Consensus	Trust vs. Pop.	Fixed Popularity	Cons. vs. Div.
All members	0.537	**All members**	0.543	All members	0.644
5 members	0.512	**5 members**	0.762	5 members	0.800
2 members	0.784	**2 members**	0.700	2 members	0.688
All interfaces	**0.597**	**All interfaces**	**0.664**	All interfaces	**0.710**

In Table XVI we report very similar D-values across conditions. In total there were 272 recordings (out of 456) for which D-values were computed, i.e. 60% of all cases amounting for 128 changes and 144 same values yielding a change rate of 47%, i.e. people changed their rating almost once at every 2 songs! Larger D-values are computed in the following conditions: “2 other members” for trust vs. popularity (0.738), consensus vs. divergence (0.729). When we fixed he divergence condition we report higher D-value again for “2 closest members” when comparing trust with popularity (0.756). The same pattern can be observed for “2 closest members” in the “consensus vs. divergence” condition. However, the greatest absolute rating differences are reported for the “5 closest members” when we have a fixed consensus and compare trust with popularity (0.700) and when we fixate popularity and compare consensus with divergence (0.800).

In summary, this data shows that people change to a greater extent their preferences when they see 2 and 5 others other members’ ratings. The extent in this case is the D-value, or the absolute difference between initial and final rating.

7.6 Individual Satisfaction and Group Size

In the following analysis we use the following 24 contributed songs without self-contributions. For each participant in the table below we extract the number of ratings, we compute the average rating corresponding to the individual satisfaction and analyze this by sets of 8 interfaces: all other members, 5 and 2 closest members.

Table XVII. Individual satisfaction statistics for sets of 8 interfaces

Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9	Total/Avg.
#ratings	15	3	13	24	18	5	22	21	11	**132**
All other members	3	3	3.7	2.9	3.1	3.7	2.1	2	2.2	**2.86**
5 closest members	3.4	4	3	3	4.2	2	2.4	1.7	1	**2.74**
2 closest members	2.8	1	3.5	3.3	3.2	3	1.9	2.9	4.3	**2.88**
All interfaces	3.1	2.7	3.5	3	3.4	3.2	2.1	2.2	2.7	**2.83**

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Total/Avg.
#ratings	5	8	12	9	6	9	8	13	17	22	**109**
All other members	4	3.5	3.8	4.3	3	4.5	2.8	3.2	4	3	**3.61**
5 closest members	5	3.8	4.5	4.5	4.3	3	2	2.3	3.7	3.4	**3.65**
2 closest members	2.5	3	4.3	3.7	2	2.5	3.5	2	2.7	2.2	**2.84**
All interfaces	3.8	3.5	4.2	4.1	3.3	3.6	2.8	2.5	3.4	2.9	**3.37**

In Group1 we report a small variation of both group and individual satisfaction per sets of 8 interfaces. For “all other members” the average group satisfaction is 2.86, for “5 closest members” 2.74 and for “2 closest members” 2.88 yielding an average satisfaction for Group1 for all 24 songs without self-contributions of 2.83.

In Group2 the situation is different: most members are satisfied when seeing “all other members” (3.61) and “5 closest members” (3.65) preferences but much less satisfied when seeing “2 closest members” preferences (2.84). However, on average, the second group is much more satisfied with others’ contributions for the 24 songs than the first group: average group satisfaction of 3.37 compared with 2.83.

7.7 Subjective and Objective Data Correlation

In order to have a deeper understanding of social influence we correlate subjective results from familiarity and trust measurements with eye-tracking data.

7.7.1 Familiarity, Trust and Eye-Gaze Correlation

First, we consider 1 to 5 familiarity and trust subjective data and normalized them by the sum of all ratings turning the rating values into 0 to 1 scores. We use this normalization process to weight more equally the familiarity and trust ratings per participant, accounting for different rating behavior. For example, if participant Pi gave only ratings of 1 to other members and Pj only ratings of 3, these will all be normalized to the same value. Since most of the ratings of our participants fall into this category (in general, subjects gave fewer familiarity and trust ratings different from 1 than minimal values) the normalization step helps us quantify subjective values better.

For example in table XVIII participant P1 submitted the values of 1, 4, 2 and 1 for familiarity and 2, 4, 3, 2 for trust for 4 other members, P2, P3, P4 and P5. The sum of all 4 submitted familiarity values is 8 (1+4+2+1). To normalize each familiarity rating to a value from 0 to 1 we divide each rating to the sum of ratings yielding 0.125 for P1, 0.5 for P2, 0.25 for P3 and 0.125 again for P5. For the values of 2, 4, 3 and 2 for trust we normalize by 11 and obtain: 0.18 for P1, 0.36 for P2, 0.28 for P4 and 0.18 for P5. Taking the average of familiarity and trust it yields a rating of 1.5 (or a score of 0.16) for P2), 4 (or 0.42) for P3, 2.5 (or 0.26) for P4 and 1.5 (or 0.16) for P5. P1 also created 6 horizontal correspondences by looking at P2, 13 for P3, 3 for P4 and 5 for P5. We divide each number by the sum (27) and obtain the normalized eye-gaze: 0.22 for P2, 0.59 for P3, 0.11 for P4 and 0.18 for P5.

Table XVIII. Example of normalizing familiarity, trust and eye-gaze for one participant

P1	P2	P3	P4	P5	Sum
Familiarity	1	4	2	1	8
Norm. Fam.	0.125	0.5	0.25	0.125	1
Trust	2	4	3	2	11
Norm. Fam.	0.18	0.36	0.28	0.18	1
Avg. Fam. & Trust	1.5	4	2.5	1.5	9.5
Norm.	0.16	0.42	0.26	0.16	1
Eye-gaze	6	13	3	5	27
Norm. Eye-gaze	0.22	0.59	0.11	0.18	1

Alongside with the above we extract the eye-tracking social data in the following way. For each of the 24 interfaces, we count the number of times each participant created vertical correspondences for one of his group members. For each interface we have a binary value: 1 meaning that the current participant Pi looked at participant Pj's ratings (where i is between 1 and 9 for the first group and between 10 and 19 for the second one and j takes all remaining values except i in the group) and 0 meaning that he did not.

This process allows us to identify the most influential group members for each participant. Next, we count the total number of correspondences for the sets of 8 interfaces corresponding to "all other members", "5 closest members" and "2 closest members". An example is given below for Group1 and total number of correspondences per participant. Out of the maximum possible number of associations, which is 24 for one other participant, P18 created 18 correspondences for P5. Thus, P8 followed the most P5, then P2 (11 associations) and then P7 (10 associations). This information allows us to identify the most influential members for each participant, eye-tracking-wise.

We measure the attention to others' rating preference with the method presented above. In order to balance the different gaze patterns among participants (e.g. P11 created a max.

of 23 of horizontal patterns for P17 alone and 95 in total whereas P13 a maximum of 3 for P15 and a total of 9 for all members) we normalize all number of correspondences in the table by the maximum per participant (this total number can be 24, if the person was among the 2 closest members, 16, if he was among the 5 closest members, but not in the 2 closest members set, or 8, otherwise) multiplied by 5 to obtain values which we will compare next with familiarity and trust.

Table XIX. Number of horizontal correspondences per participant

	P1	P2	P3	P4	P5	P6	P7	P8	P9	Max (24)	Total
P1		6	0	13	3	0	5	3	9	**13**	**39**
P2	5		4	7	11	2	5	3	3	**11**	**40**
P3	3	3		3	2	0	0	2	4	**4**	**17**
P4	0	0	0		2	1	3	0	1	**3**	**7**
P5	4	7	4	4		1	8	6	3	**8**	**37**
P6	4	6	0	3	12		10	6	6	**12**	**47**
P7	4	12	0	12	12	4		4	2	**12**	**50**
P8	4	11	0	7	18	4	10		5	**18**	**59**
P9	4	7	0	5	9	4	15	4		**15**	**48**

	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Max (24)	Total
P10		5	7	9	12	6	4	11	5	2	**12**	**61**
P11	11		7	6	3	14	6	23	10	15	**23**	**95**
P12	11	4		12	12	17	5	2	6	2	**17**	**71**
P13	0	1	1		1	3	0	1	0	2	**3**	**9**
P14	9	2	3	6		6	2	5	3	2	**9**	**38**
P15	8	0	6	4	5		1	1	0	1	**8**	**26**
P16	7	3	3	10	12	15		4	2	1	**15**	**57**
P17	9	5	3	3	11	5	3		6	8	**11**	**53**
P18	2	4	0	4	4	1	3	4		9	**9**	**31**
P19	10	12	4	4	6	4	3	16	15		**16**	**74**

The new values are reported in Table XX below. We re-scaled the 0 to 24 interval to one from 0 to 5 in order to compare subjective and objective data. Furthermore, we are interested to analyze the extent to which this normalized correspondence data represents each group accounting for the familiarity and trust values.

Table XX. Normalized horizontal correspondences per participant

	P1	P2	P3	P4	P5	P6	P7	P8	P9
P1		2.30	0	5	1.15	0	1.92	1.15	3.46
P2	2.27		1.81	3.18	5	0.90	2.27	1.36	1.36
P3	3.75	3.75		3.75	2.5	0	0	2.5	5
P4	0	0	0		3.33	1.66	5	0	1.66
P5	2.5	4.37	2.5	2.5		0.62	5	3.75	1.87
P6	1.66	2.5	0	1.25	5		4.16	2.5	2.5
P7	1.66	5	0	5	5	1.66		1.66	0.83
P8	1.11	3.05	0	1.94	5	1.11	2.77		1.38
P9	1.33	2.33	0	1.66	3	1.33	5	1.33	

	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19
P10		2.08	2.91	3.75	5	2.5	1.66	4.58	2.08	0.83
P11	2.39		1.52	1.30	0.65	3.04	1.30	5	2.17	3.26
P12	3.23	1.17		3.52	3.52	5	1.47	0.58	1.76	0.58
P13	0	1.66	1.66		1.66	5	0	1.66	0	3.33
P14	5	1.11	1.66	3.33		3.33	1.11	2.77	1.66	1.11
P15	5	0	3.75	2.5	3.12		0.62	0.62	0	0.62
P16	2.33	1	1	3.33	4	5		1.33	0.66	0.33
P17	4.09	2.27	1.36	1.36	5	2.27	1.36		2.72	3.63
P18	1.11	2.22	0	2.22	2.22	0.55	1.66	2.22		5
P19	3.12	3.75	1.25	1.25	1.87	1.25	0.93	5	4.68	

Now, we align the normalized eye-tracking correspondences alongside all familiarity and trust ratings for all participants in the same group. The correlation values for the two groups are highlighted in the table below (Table XXI): 0.509 for normalized familiarity and normalized eye-gaze in Group1 and 0.581 for Group2 and 0.528 for normalized trust and normalized eye-gaze in Group1 and 0.585 for Group2. Thus, trust proves to be slightly better correlated with normalized eye-gaze than familiarity, especially for the first group. This high values show that actually both the familiarity and trust networks correspond to a great proportion with the eye-tracking networks derived from the normalized eye-gaze data. In the following, we further improve these values by normalizing both familiarity and trust and re-computing correlation values.

We create normalized distributions of familiarity, trust and eye-gaze and correlate them for all participants in the same group (Table XXI). Finally, we compute an average correlation value for each two conditions.

Table XXI. Correlation between normalized subjective (familiarity and trust) and objective (eye-tracking) data

Condition	Group1	Group2
Norm. Fam. and Norm. Eye-gaze	0.509	0.581
Norm. Trust and Norm. Eye-gaze	0.528	0.585
Norm. Avg. Fam. and Trust and Norm. Eye-gaze	0.546	0.624

The results from table XX can be categorized into three significant findings: the first is that the high overall correlation values denote strong interconnection between what participants declared (in terms of familiarity and trust) and their actual (eye-tracking) behavior.

Secondly, correlation values between familiarity and eye-gaze (0.509 vs. 0.581), trust and eye-gaze (0.528 vs. 0.585) and the average between familiarity and trust and eye-gaze (0.546 vs. 0.624) for the second group are higher than for the first one. Indeed, the familiarity and trust data match the eye-gaze patterns better for this group than the other. Social relationships between the participants in the second group are mapped more precisely to their decision behavior than for members in the first group. This shows that participants followed more the persons they trust most (literary!). This group's higher connectivity among group members influences their decision: participants follow more closely the members they trust more and less those that they do not know. Thus, the subjects' actual behavior can be well-mapped to their subjective evaluations.

The third finding arises by comparing the rows of the table. We notice that trust is better correlated with eye-gaze than familiarity. However, highest correlation is obtained by combining the two: computing average familiarity and trust and normalize the new values we obtain highest correlation for both groups: 0.546 for Group1 and 0.624 for Group2.

To conclude, both normalized familiarity and normalized trust represent a good mapping of the normalized eye-tracking gaze. For the more connected group, the gaze behavior corresponds more to the familiarity and trust (normalized) values. Moreover, even though trust information is slightly better than familiarity for mapping subjective evaluation with eye-gaze, the average between the two represents a better indicator of participants' individual behavior.

7.7.2 Eye-tracking Correspondences

In this section we present the aggregated values for the total and average number of horizontal correspondences per participant for the two groups. As in the previous section we compute the total number of correspondences per group and per condition.

Table XXII. Number of correspondences per group and interface condition

Group1	**Total #Correspondences**	**All other members**	**5 closest members**	**2 closest members**
	377	**201**	**134**	**42**
	Avg. #Corresp. / Part.	**First 8 Interfaces**	**Second 8 Interfaces**	**Last 8 Interfaces**
	41.89	**22.33**	**14.89**	**4.67**
	Avg. #Corresp. / Part. / Interface	**First 8 Interfaces**	**Second 8 Interfaces**	**Last 8 Interfaces**
	1.75	**2.79**	**1.86**	**0.58**
Group2	**Total #Correspondences**	**First 8 Interfaces**	**Second 8 Interfaces**	**Last 8 Interfaces**
	502	**260**	**180**	**62**
	Avg. #Corresp. / Part.	**First 8 Interfaces**	**Second 8 Interfaces**	**Last 8 Interfaces**
	50.2	**28.89**	**20**	**6.89**
	Avg. #Corresp. / Part. / Interface	**First 8 Interfaces**	**Second 8 Interfaces**	**Last 8 Interfaces**
	2.09	**3.21**	**2.22**	**0.77**

In addition to the above table (Table XXII) we offer a graphical representation of the relative differences across groups and conditions. In the figure below (Figure 24) we report very similar percentages across the two groups: 53% (201 correspondences) of all number of correspondences in Group1 and 52% (260) in Group2 were formed in the "all other members" condition, 36% (134 and 180) in both groups were formed in the "5 closest members" condition and 11% (42) in Group1 and 12% (62) in Group2, respectively, were formed in the "2 closest members" final condition. These relative percentages are in accordance with the total information included in all interfaces according to the proportion: 8:5:2 for Group1 and 9:5:2 for Group2.

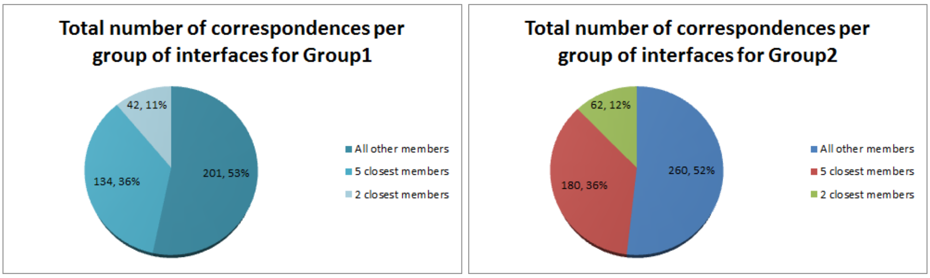


Figure 24. Total number and relative distribution of visual correspondences per group

We average the number of correspondences per participant and per interface highlighting the relative percentage differences. In Group1 the average number of correspondences per participant is 22.33 for all other members, 14.89 for 5 closest members and 4.67 for 2 closest members, whereas in Group2 higher values denoting more attention are marked as follows: 28.89 correspondences for all other members, 20 for 5 closest members and 6.89 for 2 closest members.

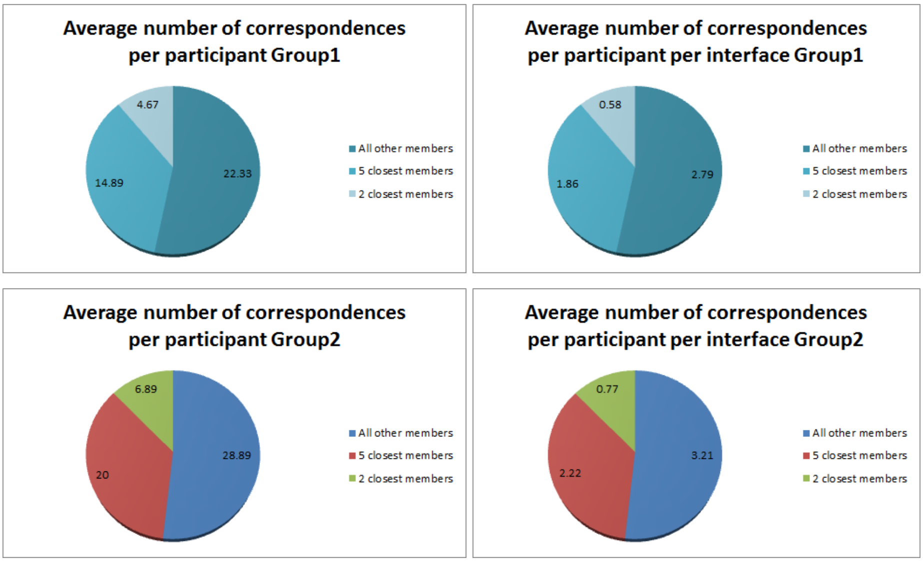


Figure 25. Number of correspondences per participant and per interface per group

Further dividing this numbers by 8 (the number of interfaces for each condition) we report that the more connected group was significantly more attentive to the eye-tracking names' and ratings as follows: 3.21 compared with 2.79 for all other members, 2.22 compared with 1.86 for 5 closest members and 0.77 compared with 0.58 for 2 closest members. Altogether, these results prove that despite the relatively balanced attention among the 3 conditions for both groups, the members of the more connected group created more gaze patterns being most attentive to the interface information.

7.7.3 Top 3 Closest Members Mapping

In the table below (Table XXIII) we compare the top 3 closest members computed using three measures: familiarity, trust and eye-tracking. The first column indicates the list of participants (Group1 at the top), the next 3 columns mark the closest 3 members as indicated by the subjective familiarity values input by the participants, the following 3 columns indicate the closest members based on trust and the last 3 columns contain the top 3 closest members based on eye-tracking horizontal correspondences: the more visual associations one subject made between other participants' names and their ratings, the closer that member is in our ordering.

As mentioned in the previous sections trust closely follows familiarity information with a very high accuracy: top 3 most trusted members are the same (100% accuracy) as top 3 based on familiarity for: P1, P2, P3, P4, P7, P9 in Group1 and P10, P12, P13, P18 and P19 in Group2. We are now interest to compute the extent to which top 3 closest members computed via music preference and eye tracking correspondences correspond to the top 3 computed via familiarity and trust values. To do so we compute the overlay values in Table XXIV.

We compute the overlay between each pair of variables from familiarity, trust, eye-tracking and music preference by comparing sets of 3 elements. Taking P6 as example we report the top 3 familiarity set as: {P5, P7, P9}, the trust set: {P5, P7, P9}, the eye-tracking set: {P5, P7, P8} and the music preference set: {P8, P3, P9}. The overlay between each two sets corresponds to the identification of each member from the 3 variables in the two collections. Comparing familiarity and trust it yields that that the two sets are identical, thus the overlay is 100%. Comparing the familiarity and eye-tracking sets P9 from the first set was replaced by P8 in the second one, thus the overlay is 66%. Comparing familiarity and music preference we report than only P9 is in both subsets: overlay of 33%. The same computation is done for trust and eye-tracking and trust and music preference.

Table XXIII. Comparing top 3 closest members based on familiarity, trust, eye-tracking and music preference

	Familiarity			Trust			Eye-tracking			Music preference		
	Fa1	**Fa2**	**Fa3**	**Tr1**	**Tr2**	**Tr3**	**ET1**	**ET2**	**ET3**	**MP1**	**MP2**	**MP3**
P1	P4			P4			P4	P9	P2	P2	P4	P7
P2	P3	P6	P7	P3	P6	P7	P5	P4	P7	P1	P4	P7
P3	P2	P5	P1	P2	P5	P1	P9	P2	P1	P8	P6	P7
P4	P1			P1			P7	P5	P6	P7	P2	P1
P5	P7	P8	P6	P7	P8	P6	P7	P2	P8	P3	P8	P1
P6	P5	P7	P9	P5	P7	P9	P5	P7	P8	P8	P3	P9
P7	P5	P9		P5	P9		P5	P4	P2	P4	P3	P2
P8				P5	P7	P9	P5	P2	P7	P3	P6	P5
P9	P7			P7			P7	P5	P2	P6	P3	P8
P10	P17	P18	P14	P17	P18	P14	P14	P17	P13	P18	P12	P17
P11							P17	P19	P15	P14	P13	P18
P12	P15	P10	P17	P15	P10	P17	P15	P13	P14	P15	P18	P10
P13	P15	P12	P17	P15	P12	P17	P15	P12	P11	P14	P18	P11
P14	P10	P15	P17	P10	P12	P17	P10	P15	P13	P11	P13	P18
P15							P10	P12	P14	P12	P18	P14
P16							P15	P14	P13	P11	P14	P17
P17							P14	P10	P19	P10	P19	P11
P18	P19			P19			P19	P17	P14	P10	P12	P13
P19	P17	P18	P11	P17	P18	P11	P17	P18	P11	P13	P11	P17

For empty values in the table above we compare only the other remaining elements (e.g. 2 instead of 3). If there are no values no computation is made.

In Table XXIV we summarize the overlay values computed for all members and averaged per group. First of all in both groups we obtain similar overlay values. Comparing familiarity and trust yields an overlay of 100% for Group1 and 94.33% for Group2; comparing familiarity and eye-tracking it yields an overlay of 66.25% for

Group1 and a higher percentage for Group2: 77.5%; comparing trust and eye-tracking we obtain an overlay of 58.88% for Group1 and about the same, 60.66% for Group2.

Table XXIV. Overlay between familiarity, trust, eye-tracking and music preference

Participant	Fam.-Trust	Fam.-ET	Trust-ET	Fam.-MP	Trust-MP	ET-MP
P1	100%	100%	100%	100%	100%	66%
P2	100%	33%	33%	33%	33%	66%
P3	100%	66%	66%	0%	0%	0%
P4	100%	66%	0%	100%	100%	100%
P5	100%	66%	66%	33%	33%	33%
P6	100%	66%	66%	33%	33%	33%
P7	100%	33%	33%	0%	0%	66%
P8			66%		33%	33%
P9	100%	100%	100%	0%	0%	0%
Average Gr.1	**100%**	**66.25%**	**58.88%**	**37.37%**	**36.88%**	**44.11%**
P10	100%	100%	66%	66%	66%	33%
P11						0%
P12	100%	33%	33%	66%	66%	33%
P13	100%	100%	66%	0%	0%	33%
P14	66%	66%	33%	0%	0%	33%
P15						66%
P16						33%
P17						66%
P18	100%	100%	100%	0%	0%	0%
P19	100%	66%	66%	33%	33%	66%
Average Gr.2	**94.33%**	**77.50%**	**60.66%**	**27.50%**	**27.50%**	**36.30%**
Both groups	**97.57%**	**71.07%**	**59.60%**	**33.14%**	**33.13%**	**40.00%**

For music preference we notice a descending trend. Generally, none of: familiarity, trust or eye-tracking data are useful to predict participants' music preference with high accuracy. We report an overlay between familiarity and music preference of 37.37% for Group1 and only 27.50% for Group2, about the same values for the overlay between trust and music preference: 36.88% for Group1 and 27.50 for Group2 and slightly bigger values for the overlay between eye-tracking and music preference: 44.11% for Group1 and 36.30% for Group2. For the two groups, we conclude using the last row that there is a high overlay between the first 3 variables: familiarity and trust (97.57%), familiarity and eye-tracking (71.07%) and trust and eye-tracking (59.60%) but low overlay between the three variables and music preference taken separately: 33.14% for familiarity and music preference, 33.13% for trust and music preference and 40.00% for eye-tracking and music preference.

To conclude this section, the results include above confirm that individuals trust more the people they are most familiar with: an overlay between the two variables is 97.57%! In addition to this, the extracted eye-tracking objective data supports the previous observation resulting in a high overlay between familiarity and eye-tracking (71.07%) and trust and eye-tracking (59.60%), respectively.

For certain participants who submitted normalized ratings for familiarity and trust the eye-tracking accuracy predicting one's top 3 closest members was as high as 100%. However, the main limitation of this type of computation is that it relies on incomplete data, more precisely; there are no differences across sets of familiarity and trust evaluations submitted by some participants: it is not possible to differentiate clearly among the top3 members based on especially familiarity and trust. Eye-tracking and music preference provide more accurate identification of top 3 candidates.

7.8 Other Members' Influence

In this section we analyze the impact of a specific group of members on each participant's decision. We start by highlighting the process of extracting horizontal correspondences per participant by considering an example.

7.8.1 Analysis Process

In order to objectively measure the group influence with respect to the stimulus information presented to them, represented by the songs' name and artist (AOIS) and other members' names (AOIU) together with their ratings (AOIR), for each of our participants in the two groups we extracted the visual horizontal correspondences with respect to each other member's preferences. In the case such a correspondence was detected we marked the respective participant's name together with his rating. In the case there was no such correspondence detected we would assume that the respective person did not exert any influence at all on the decision of the current participant watching the respective stimulus. In short, if a person did not look at another person's name and rating we consider that there was no decision influence.

As an example, we consider participant P8 from Group1. For the first set of 8 interfaces displaying all other 8 members' names and ratings, he looked only at P1, P3, P4 and P7 and their ratings for interface 1, P1 and P2 for interface 2, P2, P3, P4, P5 and P6 for interface 3, etc.

For the second group of 8 interfaces displaying 5 closest members' preferences we apply the same extraction rule: for interface 9 P8 looked at P1, P3 and P4, for interface 10, P8 looked only at P2, etc. Finally, for the last 8 interfaces displaying only 2 closest members we extract the horizontal correspondences as: P8 looked at P1 and P4 for interface 17, P8 looked at only P1 for interface 18, P8 looked at only P4 for interface 19, etc.

In this way we filter out irrelevant information selecting only the correspondences that had an impact for each participant's decision. However, up to this point in our presentation, we do not discern between low and high impact since we do not know whose rating's matter the most to the current user.

7.8.2 Rating Strategy

We carry on an in-depth analysis to compute the correlation between the rating decision stated after the group influence stimulus and the information displayed in the social interfaces. Here we are interested to test the rating strategy, on average, of both groups whether it corresponds to a minimum, maximum or average of all the ratings displayed in the 24 interfaces.

The analyzed data is in the form of members' names and their ratings which can include none, all, or any number of members per interface. Thus, we compute the minimum, average and maximum of the ratings in the interface in order to see which strategy people follow when stating their own evaluations. Our hypothesis is that most participants will select the average decision of the group from which they would deviate only slightly.

To compute the correlation between each participant's rating and the minimum, average and maximum, respectively, of the interface values we proceed as follows. First of all, for all 24 interfaces we list each participant's rating values alongside with the minimum, maximum and average group satisfaction of only the members that he looked at from the interface - not all of the ratings displayed, but only those that matters specifically for each participant! In this way we perform a "personalized" analysis aiming to understand individual judgment according to the information that was filtered individually.

By removing the ratings that were not observed, thus the not-relevant ones, we restrict our analysis to the factors which indeed influenced our subjects. We additionally, add the number of horizontal associations that each participant made per interface. In the table below we include the example of P8 for half of the interfaces (the first 12).

From Table XXV below, P8 gave the following vector of ratings: {2, 4, 4, 1, 3, 1, 4, 2, 1, 2, 2, and 3} to each of the first 12 songs that he listened to. As a member of Group1 he saw all other members' (8 in total) preferences in the first 8 interfaces. He looked at 4, 5, 7, 3, etc. members' preferences displayed to him. For all of this we used the interface design template and each individual rating in order to compute the average, minimum and maximum of the respective ratings that the participant gazed at. As such, for the first interface, the 4 members' preferences that P8 looked at have an average of 3.75 with a minimum of 3 and a maximum of 5.

For the next interface, P8 looked at 5 members' preferences having an average of 4.6, a minimum of 4 and a maximum of 5. For this interface he submitted a rating of 4. The same process applies to the other interfaces.

Table XXV. Example of judgment analysis

Interface	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Rating	2	4	4	1	3	1	4	2	1	2	2	3
#Corresp.	**4**	**5**	**7**	**3**	**5**	**3**	**3**	**3**	**4**	**3**	**4**	**4**
Average	3.75	4.6	4	3.66	3.4	4	4.66	3.66	4.5	3.66	3.4	4.33
Minimum	3	4	3	3	2	3	4	3	4	3	3	4
Maximum	5	5	5	5	5	5	5	5	5	5	5	5

More generally, the main results in this section are based on the following data:

- A = Minimum of all ratings that participants looked at by creating clear visual horizontal correspondences with other people's names in the interface;
- B = Average of all ratings data as described in the example above;
- C = Maximum of all ratings data as described in the example above;
- D = Each participant's decision data (ratings) after seeing each of the 24 interfaces, e.g. P8.

The number of correspondences is used for the purpose of calculating the values of B, C and D for each subject. It is interesting to note that out of the total number of correspondences in Group1 (522) participants focused on as many as 52.83% (Table XXVI). More specific, they at 44.25% out of the total number of correspondences for the interface displaying "all other members" preferences, for 47.50% out of all possible for "5 closest members" and as much as 66.75% for "2 closest members". For the more connected group, Group2, we report enhanced percentages especially for the "all other members" (45.87%) and "5 closest members" (58.25%) resulting in 57% of attention gazes recorded out of a maximum of 669 possible.

Table XXVI. Eye-gaze and attention statistics

	Total # Corresp.	First 8 interfaces	Second 8 interfaces	Last 8 interfaces
Group1	**522**	**8**	**5**	**2**
Percent from total possible	52.83%	44.25%	47.50%	66.75%
Group2 #Corresp.	**669**	**9**	**5**	**2**
Percent from total possible	57.00%	45.87%	58.25%	66.87%

We correlate D with each of the computed eye-tracking data (A, B and C) and average the correlation values per group. The values in the table below (Table XXVII) show unexpected results comparing Group1 with Group2. Correlation values for the first group are much higher than those for the second group suggesting that people in this group rated closest to the average of group decision saw in the interfaces (39.71%), then to the minimum (33.13%) and then to the maximum (27.44%). In the second group the correlation values are close to zero: members in this group would not state their preferences based on the minimum, average or maximum of all other members they looked at. Rather, they would use a different rating strategy.

Table XXVII. Correlation values between each participant's ratings and the minimum, average and maximum of all others' ratings

	D&A	D&B	D&C
Group1	33.13%	39.71%	27.44%
Group2	14.70%	9.14%	-4.46%
Average	**23.91%**	**24.42%**	**11.49%**

For more detail we highlight the correlation values between each participant's rating and the minimum, average and maximum of all the ratings he gazed at.

Table XXVIII. Correlation values between group ratings and minimum, average and maximum of observed ratings

	MIN	AVG	MAX
P1	-0.49	-0.42	-0.34
P2	0.11	0.04	0.04
P3	0.20	0.28	0.25
P4	0.74	0.78	0.66
P5	0.45	0.44	0.42
P6	0.63	0.71	0.60
P7	0.22	0.44	0.41
P8	0.29	0.34	0.36
P9	-0.14	-0.22	-0.03
Average	**0.22**	**0.26**	**0.26**

	MIN	AVG	MAX
P10	0.35	0.52	0.44
P11	0.89	0.92	0.86
P12	0.37	0.38	0.14
P13	0.65	0.84	0.73
P14	0.80	0.74	0.47
P15	0.73	0.86	0.75
P16	0.65	0.62	0.46
P17	0.60	0.68	0.48
P18	0.75	0.84	0.76
P19	0.83	0.91	0.83
Average	**0.66**	**0.73**	**0.59**

We obtain rather small correlation values (Table XXVIII top) with the minimum: 0.22 for Group1 and 0.66 for Group2, average: 0.26 for Group1 and 0.73 for Group2 and maximum: 0.26 for Group1 and 0.59 for Group2. If we eliminate P1 from the analysis due to very few ratings and P9 due to negative correlation between his rating behavior and the ratings from the interfaces we obtain a correlation of 0.38 between the reported group ratings and the minimum in the interfaces, 0.43 between the average and the group ratings and 0.39 between the maximum and the group ratings for Group1.

Even with these “improved” values for Group1 we obtain larger correlation values for the more connected group (Table XXVIII – bottom). This proves once again that the members of the second group “obeyed” more the rating information displayed on the interfaces. They were more inclined to accept these music recommendations and to align their preferences to those saw on the interfaces.

Furthermore, this proves the highest alignment with respect to the average condition. In the same time, the correlation values of re-submitted ratings and minimum and maximum, respectively, are only slightly smaller.

7.8.3 2 Closest Members’ Influence

We see the above result encouraging for looking for more details about other characteristics of each group. If all other members’ ratings do not seem to matter qualitatively much for the more social group, then by looking at 2 closest members we would be able to understand better how social relationships actually dictate the rating change.

The question we try to respond to here is: “How much influence do 2 closest members exert on one’s own decision?”

For this we narrow down our data to extracting only the ratings of 2 closest members, per participant, throughout the 24 interfaces. Continuing the example from the previous section we identified that P1 and P4 are the closest members of P8 (based on the familiarity and trust networks discussed in the beginning).

We put together the following data for our analysis:

- 2 closest member ratings, e.g. P1 and P4's ratings ;
- A' = Minimum of 2 closest members ratings, e.g. minimum(P1, P4);
- B' = Average of 2 closest members ratings, e.g. average(P1, P4);
- C' = Maximum of 2 closest members ratings, e.g. maximum(P1, P4);
- D = Each participant's decision data (ratings) after seeing each of the 24 interfaces, e.g. P8.

Table XXIX. Correlation values between each participant's ratings and the minimum, average and maximum of 2 closest members' ratings

	D&A'	D&B'	D&C'
Group1	22.51%	26.75%	26.52%
Group2	66.72%	73.55%	59.62%
Average	**44.61%**	**51.15%**	**43.07%**

The correlation values between each participant's ratings and the minimum, average and maximum of 2 closest members' ratings that he gazed at are presented in Table XXIX. The results summarized in the table above show a reversed situation from the previous table.

Indeed, members in the second group align more to 2 closest members' preferences. We notice the highest correlation, 73.55% for the average value of the 2 ratings, followed by the minimum (66.72%) and the maximum (59.62%). Overall, these results imply that members in the second group follow to large extent 2 closest members' preferences. By opposition, for Group1, results do not show a high correlation between each participant's ratings and the ratings of his 2 closest peers, 26.75% for average, 26.52% for the maximum of ratings and 22.51% for the minimum of ratings.

As we calculated the correlation values for each participant we report 2 negative values for only 2 participants in the first group: P1 and P9. Without these negative values the correlation values for the remaining 7 members in Group1 would increase with more than 10% on average: 43.69% for average, 39.50% for maximum and 38.19% for minimum. Furthermore, in this group the majority (all other group members) have a stronger impact for individual decision rather than these 2 closest members. This fact is proved by comparing the second row in table 15 with the second row in table 16. In Group2, however, social influence is most prominent for 2 closest members compared with all other members. For this group the influence of all others is least whereas people align their preferences most (highest correlation values) with 2 of their closest peers.

Altogether the results in this section prove the high impact that 2 closest members' have on one's decision for a connected group. If sparse relationships are present between group members then each subject tends to align his decision more to the majority rather than to only few closest members. On the other hand, given strong social bounds between group members, participants are more influenced by the decision of few of their closest members when making their own judgment obtaining very high correlation values, as much as 92.83% for P11, 91.02% for P19, 86.02% for P15, 84.62% for P13, and 84.57% for P17 – all for the average case. Between minimum, average and maximum, participants align most to the average of the ratings displayed to them in the vast majority of cases.

7.9 Symmetry and Transitivity

Symmetry refers to the following question: "if Pi often looks at Pj's preferences is it true that Pj also looks at Pi's ratings?" To formulate our answer we first need to look at the familiarity and trust matrix per group. Our data shows that, in general, the familiarity and trust ratings are reciprocal but at different intensities. For instance it was commonly observed that one's closest familiar / trusted members also submit similar high familiarity / trust evaluations in return.

As example, we consider P2 who is familiar mostly with P3, P6 and P7 and trusts most the same participants. The eye-tracking data shows P2 following most P5 (11 interfaces), P4 (7 interfaces), P1 (5 interfaces) and P7 (5 interfaces). For symmetry we need pairs of individuals which have similar familiarity and trust evaluations among themselves and others. There are only few such cases. For example, P18 submitted a familiarity rating of 2 (highest) for P19 and P19 submitted 4 (again the highest) whereas trust values indicate reciprocity. The eye-tracking data shows 9 gazes of P18 to P19 (highest number) and 15 of P19 to P18 (second highest).

Analyzing transitivity is a rather complex issue since it involves decision and eye-tracking models which are difficult to quantify and interpret. However, at a higher level, our data supports reciprocal relationships for a wider group: people follow a small group of usually 3 to 4 persons they are familiar with and trust most – more than 80% of eye-tracking correspondences are based on closest group members.

7.10 Discussion

This study represents a new approach to the study of group influence by incorporating objective eye-tracking data alongside users' subjective perceptions of familiarity and trust towards other members.

It stands out with respect to other advancements in the study of group influence through the following key strengths:

(1) It compares subjective evaluation with objective data measurements;
(2) It uses an extensive dataset of eye-tracking fixation points;
(3) It is based on group selected preference instead of external data;
(4) It defines a group influence index based on group alignment useful for understanding individual preference variations and alignment;
(5) It compares individual and group satisfaction;
(6) It outputs change rate values and decision alignment.

7.10.1 Advantages

One of the key advantages of our study is related to the experiment design. First of all we found 2 relatively large groups of people that knew and trusted others to a somehow great extent. This familiarity and trust aspect allowed us to develop an in-depth analysis based on significant fixation times and important eye-gaze correspondences. The first group included master and PhD students attending the same university course whereas the second group included PhD and Post-Docs than work together on related research and belong to neighboring labs.

Having an extensive rating and eye-gaze data-set we derived significant findings proving group alignment. Some of the most important findings rely on the relevance of our data after statistical testing. In the same time, we benefit from the eye-tracker's precision in identifying all useful associations necessary for our change rate analysis.

Another very strong determinant of our approach is that of creating permanent connections between subjective and objective data, allowing us to build an "eye-tracking" social network and compute correlation values to test the accuracy of predicting 3 closest members via people's eye gaze.

Last but not least, one key advantage of our research is due to identifying the rating strategy that members in each group are most inclined to adopt from minimum, average and maximum. Here we highlighted larger correlation values for the average strategy, especially for the more connected group.

In general, through our between-groups user study, we prove that the more connected group benefit from more members' attention through eye gaze fixation times and group rating alignment. Moreover, members in the more connected group deviate more from pre-stated decision and change their preferences slightly more.

7.10.2 Limitations

Main limitations are due to the relatively small number of participants, the interfaces presented to them and the eye-tracker's precision with respect to unique horizontal correspondences. We definitely need to extend our study for considering a larger segment of user-group population and, perhaps, display many more interfaces collecting longer fixation times and more horizontal correspondences.

Another condition is that for some songs, users did not spend enough time to analyze the interface but they proceeded to state a rating immediately after hearing the song. Here we could have imposed a minimal time limit before decision. However, we did not want to take into account artificially-created gaze patterns that would worsen our results.

In other cases users simply liked the songs played to them on the iPod and browsed only for curiosity their peers' ratings explicitly mentioning that they were not influenced at all by them but by the quality of music.

One important concern regarding the employed methodology might refer to the fact that we considered each songs uploaded by our participants to be equally important for the overall experiment considering participants' behavior. We are aware that due to popularity some of the songs might be liked by an overwhelming majority of people whereas others might be totally inappropriate for a party. We minimized this effect through our selection of the 24 songs out of the 39 and, respectively 47. Those songs were selected by popularity, thus minimizing the influence of cultural factors on people's tastes. However, through the experiment design we mark the importance of having self-contributed songs that subjects identify with. This increases the trust and confidence in the experiment as well as in the others' (presumed) submitted ratings.

An extension of our experiment which would include an extra dimension measuring memory would be useful to test the extent to which people change their opinions regarding self-ratings under no influence. In our current experiment we do not take into account the temporal factor and suppose that the change is only due to the social factors included through the group interfaces. In general, we expect the change which is due to the temporal factor to be much smaller than the one due to the social influence factor.

7.10.3 Using the Eye-Tracker

Overall, our study participants enjoyed listening to the 24 songs played on the iPod as well as standing still for around 20-25 minutes in front of the eye-tracker performing the experiment. Besides the initial calibration step in the beginning, users were very anxious to start the experiment and discover the information in each new interface. Despite the fact that some songs received more attention than others all participants mentioned that they enjoyed the selection of music made available to them which they would listen alone or in a group while driving, cycling, jogging, etc.

7.10.4 Participants' Comments

Our study participants enjoyed discovering other member's preferences in both the individual and group conditions. Some of them explicitly stated that they recognized who from the group contributed to the songs they rated using the eye-tracker being sure of that person's music preferences. Given the rating information available to them in the eye-tracking phase, participants mentioned that they spent a considerable amount of time gazing at others' ratings before deciding their own rating. They mentioned that the most useful interfaces are those in which relatively little information is presented, such as the ratings and names of 2 or maximum 5 closest members. Moreover, they confirmed our findings that they align more when there is an overall consensus for a given songs rather than when the standard deviation of ratings is high. In such case they felt that they could "trust" the group most. They were also tempted not to be influenced by any of the ratings and follow more their own judgment when shown all other members' preferences – this was confirmed based on the change rate values in the results section.

Comments by our participants can be grouped into 4 categories: songs, music similarity, discovery and other members. Through our open discussions all participants appreciated the connection between music and friends for discovering new songs that they might like. Selecting common songs in a group they noticed to be a great idea for a party or a common event to attend together.

Below we quote the most interesting comments we received.

- Songs:

P3 - *"I used to think that I listen to strange music that people don't like. But now I can see that my friends do like it and this makes me happy and proud."*

P8 – *"My music is unique and I used to believe that it's only me who like it. So I was not very interested in sharing it. Now I see that my friends become addicted. I will upload more."*

P15 - *"My experience tells me that I listen to various kinds of rock. Songs in GroupFun I wouldn't have expected to be that great. People really uploaded cool songs."*

P16 - *"So much new and diverse music! I didn't know most of the songs and I find them quite interesting."*

- Music similarity:

P5 – *"For a long time I wanted to share my music with my friends and did not know how to do that. Now I can see that they share the same tastes with me and it feels great."*

P7 – *"I spend some daily time on Facebook browsing information. Now, with GroupFun, music activity has become more challenging. I like receiving and giving scores to songs."*

P14 - *"I do not use Facebook at all. But the social connector in GroupFun is very effective in keeping friends together. I love listening to other people's songs: they produce excellent recommendations."*

- Discovery:

P1 - *"Very interesting songs. I didn't know that I could discover so many great songs. I will definitely listen to the GroupFun playlist often."*

P2 - *"I love rock music. And I can see that I am not the only one in GroupFun. This application helps me discover similar tastes with my friends."*

P4 - *"Almost never I hear to music in a public place. But the GroupFun idea is fantastic. Now I can see my friends' preferences and listen to them. It's so much good music here!"*

P9 - *"Sometimes I like listening to Indian music and top billboard songs do not interest me that much. I can see that my friends have very strong music preferences which are unique."*

P10 - *"Discovering my friends' music preferences is great. GroupFun really motivates me to upload more songs and interact with them more often."*

P11 - *"Even though I do not listen to music that much I find it interesting to discover other people's tastes. Some songs are really good and I would like to listen to them more often."*

- Friends:

P12 - *"I have all of my close friends on Facebook and we discuss many things on the platform. With GroupFun music enables us to get in contact even more frequently and listen to common music."*

P13 - *"GroupFun is a great platform for music. Not only that I have my own collection of songs I can listen to all the time but I can see my friends' participation."*

P19 – *"I don't like Facebook but I do like the concept of having a common playlist with my friends. GroupFun is very motivating for me to contribute seeing my friends' ratings."*

- Other:

P6 – *"I like GroupFun to keep my music preferences and allow me to access my songs from any device having Facebook."*

P17 – *“I listen to GroupFun songs even when I work. It's great not knowing what comes next. In the same time I trust my friends for these good suggestions.”*
P18 – *“This music is great! I love listening to most the songs. I find them quite interesting. The GroupFun playlist is very rich.”*

Overall, all participants highly enjoyed taking part in the experiment and found it quite a unique adventure to discover their peers’ music tastes. As strongly motivated, they involved themselves in pushing “the best” songs to the top of their group selection for the party they imagined to organize.

People are more willing to be influenced when they do not know a certain music item and are hesitant with respect to a specific rating to submit. In such case they spent more time looking at other’s ratings and confirming with the overall group decision.

Participants mentioned that since most of the songs were new to them they were curious to check their friends’ ratings. Then they did not deviate much from the ratings suggested to them in all 3 cases: all members, 5 closest members or 2 closest members.

In the first group users were more inclined to listen to a specific music genre but agreed that most of the songs were of very good taste matching their reunion. In the second group participants uploaded more songs specific from their region rather than international hits. This also led to high group agreement and decision acceptance.

Overall, we received interesting input with respect to the overall methodology, system design and experiment setup. One important aspect of the experiment was to learn about the main decision difficulties that the participants encountered and how they adapted on the go through alignment – personal judgment.

8. Conclusions and Future Research

This study represents a new approach to the study of group influence by incorporating objective both subjective (familiarity and trust) and objective (eye-tracking) data. In this report we presented our experiment design and results for understanding group influence in a music context. With the use of GroupFun, our music recommender system, we logged participants' individual ratings with no external influence and compared them with (possibly) changed preferences in a group setting. Given the scores of their peers in the same group, subjects were asked to re-submit new rating values. This between-group design facilitated in-depth discussions and analysis about group alignment, individual satisfaction and rating strategy driven by group connectivity.

We analyzed both rating and eye-tracking data. By comparing the subjective values from the familiarity and trust networks with the objective ones recorded with our tracking device we identified that the more connected or social a group is, the more correspondences subjects produce between users' names and their ratings. Users create these associations visually in order to infer individual decisions.

We identified clear visual patterns by clustering our group-influence interfaces into 3 areas of interest (AOIs) and identifying inter and intra AOI horizontal and vertical correspondences. Our eye-tracking results show that participants which felt more familiar with other members' music tastes and trusted them more looked for longer time and produced eye-tracking associations for all members in the more distant group (Group1) and for 2 and 5 closest members in the more connected group (Group2).

Furthermore, the eye-tracking data confirms the social relationships among participants: individuals followed their closest members' preferences during the eye-tracking experiment as subjectively evaluated through familiarity and trust questionnaires. High correlation values between familiarity, trust and eye-gaze support this result. By opposition, the music preference network can be interfered by neither familiarity / trust nor eye-tracking information. We identified top 3 closest members based on various approaches using the 4 parameters and computed an overlay between them.

With the use of the Pearson correlation score we showed how rating alikeness corresponds to familiarity and trust. Another innovative result present in this study is the computation of an individual group influence score which allow researchers to understand the extent to which people adapt to group changes and how much they trade-off their individual evaluation to achieve a group outcome.

The main results of our experiment point out to the fact that individuals do not have innate preferences but adapt their choices to those of the group. We computed social influence scores per participant identifying individuals who are more group-oriented and change more from those who are more self-oriented and vary their preferences less. We also calculated absolute rating differences emphasizing on the degree of change that subjects produced when deviating from initial preferences. Most importantly, we extracted the eye-tracking ratings of those participants' that each subject gazed at the most. This was useful to for two reasons: (1) first, to capture the amount of information from the interface gaze that our participants' found useful for their decision making and (2) secondly, to identify a simple strategy that subjects use, or adhere closest to, after seeing certain ratings.

Overall, our results show that the group is more important than the song collection. Despite people's inclination for specific genres of music and favorite artists in our experiment framework we proved that it is actually the group which dictates individual satisfaction rather that the music items themselves. By drawing an in-depth analysis on social influence factors through the study of subjective (familiarity and trust information) and objective (eye-gaze data) factors we proved that the higher the connectivity among group members the more satisfied individuals are and the more they adapt to the group decision. In the eyes of our participants, peer ratings are more important that the perceived quality of the songs them-selves. Furthermore, in this paper we proved that participants align their behavior to that of the group through a series of conditions: trust vs. popularity, diversity vs. consensus and across 3 types of interfaces: "all other members", "5 closest members" and "2 closest members". We identified the average strategy to correspond most to individuals rating decision with great differences between the two groups: the more connected group proves both greater attention and greater alignment with respect to the significant recorded eye-tracking correspondences.

Fur our future work we plan to:

(1) Produce an in-depth qualitative analysis of rating changes by computing correlation values for all 4 conditions: controversial vs. non-controversial and familiarity vs. rating ordering for more GroupFun users;

(2) Compare intra vs. inter AOI correspondences and interpret horizontal and vertical lines per user to study group influence and identify personality traits at an individual level rather than at a group level. We will further compute image histogram and overlay multiple interfaces from connected users to understand in more detail the extent to which users adapt to others' ratings;

(3) Draw correlations between group influence, participation and involvement identifying isolated users who are most "distant" from group decision. Further analysis is needed to compute eye-tracking-based metrics and interpret individual alignment to group preference by comparing the ratings displayed in the interfaces with those selected by each participant.

Another interesting research direction to pursue is to compare the explicit and self-revealed familiarity and trust values with implicit ones logged by GroupFun. For this we would need more member-to-member interaction by logging the information displayed and visualized each time by all participants accessing the system.

For longer term research we also plan to include in our analysis the effect of memory by testing if participants maintain a consistent evaluation of songs without social influence factors. It would be necessary to test the effect of memory on music preference at different scales: one week, two weeks, one month, two months, etc. In addition to the above, we plan to extend our experiment to include other groups of participants with the same or different number of participants. In order to generalize our results we need more eye-tracking and social evaluation information in both small and large groups. For this we plan to recruit small (3-4 members), medium (7-8 members) and large (9-10 members) groups.

In conclusion, this experiment offers a detailed understanding of some of the most important issues which need to be considered when evaluating acceptance of group decisions, such as dynamics, social relationships and relevant decision information.

9. Acknowledgements

This work was supported by SNSF grant 200021-119955. We are very grateful to all the subjects that took part in our experiment. Their comments and inputs were extremely valuable. We would also like to thank to the numerous GroupFun users who provided generous improvements suggestions along the development phase. Other constructive feedback which was highly appreciated was received during conference presentations and poster discussions.

10. References

(1) Agotnes, T., van der Hoek, W., Wooldridge, M. *Reasoning about Judgment and Preference Aggregation*. In Proceedings of the 6th International Joint Conference on Autonomous Agents and Multi-agent systems. ACM Press (2007).

(2) Baskin, J. P., Krishnamurthi, S. *Preference Aggregation in Group Recommender Systems for Committee Decision-Making*. In Proceedings of the 3rd ACM Conference on Recommender Systems. Pages 337-340. ACM Press (2009).

(3) Birkett, S., Galpin, A., Cassidy, S., Marrow, L., Norgate, S. *How revealing are eye-movements for understanding web engagement in young children*. In Proceedings of the Annual Conference Extended Abstracts on Human Factors in Computing Systems. Pages 2251-2256. ACM New York (2011).

(4) Bonhard, P., Harries, C., McCarthy, J., Sasse, M. A. *Accounting for Taste: Using Profile Similarity to Improve Recommender Systems*. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems. Pages 1057-1066. ACM Press (2006).

(5) Bu, J., Tan, S., Chen, C., Wang, C., Wu, H., Zhang, L., He, X. *Music Recommendation by Unified hyper graph: Combining Social Media Information and Music Content*. In Proceedings Multimedia. ACM Press (2010).

(6) Cui, P., Wang, F., Ou, M., Yang, S., Sun, L. *Who Should Share What? Item-level Social Influence Prediction for Users and Posts Ranking*. In Proceedings of the 34th ACM SIGIR International Conference on Research and Development in Informational Retrieval. Pages 185-194. ACM Press (2011).

(7) Castagnos, S., Jones, N., Pu, P. *Recommenders' Influence on Buyers' Decision Process*. In Proceedings of the 3rd ACM conference on Recommender Systems. Pages 361-364. ACM Press (2009).

(8) Castagnos, S., Jones, N., and Pu, P. *Eye-tracking Product Recommenders' Usage*. In Proceedings of the 4th ACM Conference on Recommender Systems. Pages 29-36. Barcelona, Spain (2010).

(9) Castagnos, S., Pu, P. *Consumer Decision Patterns Through Eye Gaze Analysis*. In Proceedings of the International Workshop on Eye Gaze in Intelligent Human Machine Interaction in conjunction with the 2010 International Conference on Intelligent User Interfaces. Pages 78-85. Hong Kong (2010).

(10) Celma, O., Cano, P. *From Hits to Niches? or how Popular Artists can Bias Music Recommendation and Discovery*. In Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition. ACM Press (2008).

(11) Chen, H.-C., Chen, L. P. *A Music Recommendation System Based on Music Data Grouping and User Interests*. In Proceedings of the 10th International Conference on Information and Knowledge Management. Pages 231 – 238. ACM Press (2001).

(12) Chen, L., Pu, P. *Users' Eye Gaze Pattern in Organization-based Recommender Interfaces*. In Proceedings of the 15th International Conference on Intelligent User Interfaces. Pages 311-314. Palo Alto, California, USA (2011).

(13) Chen, L., Pu, P. *Eye-Tracking Study of User Behavior in Recommender Interfaces*. In: P. De Bra, A. Kobsa, and D. Chin (Eds.): UMAP 2010, LNCS 6075. Pages 375-380, Hawaii, USA (2010).

(14) Chen, Y., Harper, F.M., Konstan, J., and Li, S.X., *Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens*. American Economic Review 100, 4. Pages 1358-1398 (2010).

(15) Cialdini, R.B., Goldstein, N.J. *Social Influence: Compliance and conformity*. Annual Review of Psychology 55. Pages 591-621 (2004).

(16) Cialdini, R.B., Trost, M.R. *Social Influence: Social Norms, Conformity, and Compliance*. In The Handbook of Social Psychology, D.T. Gilbert, Fiske, S.T., Lindzey, G. (Eds.). Pages 151-192. Oxford Univ. Press (1998).

(17) Drusch, G., Bastien, J.M.C., Dinet, J. *From Gaze Plots to Eye Fixation Patterns using a Clustering Method based on Hausdorff Distances*. In Proceedings of the 23rd French Speaking Conference on Human-Computer Interaction. ACM Press (2011).

(18) Dumais, S.T., Buscher, G., Cutrell, E. *Individual Differences in Gaze Patterns for Web Search*. In Proceedings of the 3rd Symposium on Information Interaction in Context. Pages 185-194. ACM Press (2010).

(19) Fleishman, J.A. *The Effects of Decision Framing and Others' Behavior on Cooperation in a Social Dilemma*. In Journal of Conflict Resolution 32, 1. Pages 162-180 (1998).

(20) Gartrell, M., Xing, X., Lv, Q., Beach, A., Han, R., Mishra, S., Saeda, K. *Enhancing Group Recommendation by Incorporating Social Relationship Interactions*. In

Proceedings of the 16th ACM international conference on Supporting Group Work. Pages 97-106. ACM Press (2010).

(21) Grossi, D. *Unifying Preference and Judgment Aggregation*. In Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems, 1. Pages 217-224. ACM Press (2009).

(22) Kang, H., Getoor, L., Singh, L. *Visual Analysis of Dynamic Group Membership in Temporal Social Networks*. In Proceedings of ACM SIG Conference on Knowledge Discovery and Data Mining. Pages 13-21. ACM Press (2007).

(23) Masthoff, J. *The Pursuit of Satisfaction: Affective State in Group Recommender Systems*. Computer Science, Vol. 3538. Pages 297-306 (2005).

(24) Nolan, J.M., Schultz, P., Ciladini, R.B., Goldstein, N.J., Griskevicius, V. *Normative Social Influence is Underdetected*. Personality and Social Psychology Bulletin 34, 7, Pages 913-923 (2008).

(25) Popescu, G., Pu, P. *Probabilistic Game Theoretic Algorithms for Group Recommender Systems*. In Proceedings of the Workshop on Music Recommendation and Discovery in conjunction with the 5th ACM Conference on Recommender Systems. ACM Press (2011).

(26) Popescu, G., Pu, P. *What's the Best Music You Have? Designing Music Recommendation for Group Enjoyment in GroupFun*. In Proceedings of the ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems. Pages 1673-1678. ACM Press (2012).

(27) Popescu, G., Pu, P. *A Different View on Group Analysis - (Eye-) Tracking Social Influence*. EPFL Technical Report. Lausanne, VD, Switzerland (2012).

(28) Postmes, T., Spears, R., Sakhel, K., De Groot, D. *Social Influence in Computer-Mediated Communication: the Effects of Anonymity on Group Behavior*. Personality and Social Psychology Bulletin 27, 10. Pages 1243-1254 (2001).

(29) Prada, R., Otero, N., Paiva, A. *The User in the Group: Evaluating the Effects of Autonomous Group Dynamics*. In Proceedings of the International Conference on Advances in Computer Entertainment Technology. Pages 25–32. ACM Press (2007).

(30) Pretorius, M., Gelderblom, H., Chimbo, B. *Using Eye Tracking to Compare how Adults and Children Learn to use an Unfamiliar Computer Game*. In Proceedings of the 2010 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists. Pages 275-283. ACM Press (2010).

(31) Pu, P., Chen, L., Hu, R. *A User-Centric Evaluation Framework for Recommender Systems*. In Proceedings of the 5th ACM Conference on Recommender Systems. Pages 157-164. Chicago, IL, USA. ACM Press (2011).

(32) Purpura, S., Schwanda, V., Williams, K., Stubler, W., Sengers, P. *Fit4Life: The Design of a Persuasive Technology Promoting Healthy Behavior and Ideal Weight*. In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems. Pages 423-432. Vancouver, BC, Canada. ACM Press (2011).

(33) Sassenberg, K., *Common Bond and Common Identity Groups on the Internet: Attachment and Normative Behavior in On-Topic and Off-Topic Chats*. Group Dynamics 6, 1. Pages 27-37 (2002).

(34) Sukumaran, A., Vezich, S., McHugh, M., Nass, C. *Normative Influences on Thoughtful Online Participation*. In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems. Pages 3401-3410, Vancouver, BC, Canada. ACM Press (2011).

(35) Tang, J., Sun, J., Wang, C., Yang, Z. *Social Influence Analysis in Large-Scale Networks*. In Proceedings of the 15th ACM International Conference on Knowledge Discovery and Data Mining. Pages 807-816. ACM Press (2009).

(36) Turner, J.C. *Social Influence*. Wadsworth (1991).

(37) Van Maanen, P.-P., Klos, T., van Dongen, K. *Aiding Human Reliance Decision Making Using Computational Models of Trust*. In Proceedings of the ACM International Conference on Web Intelligence and Intelligent Agent Technology. Pages 372-376. ACM Press (2007).

(38) Zhu H., Huberman, B., Luon, Y. *To Switch or not to Switch: Understanding Social Influence in Online Choices*. In Proceedings of the ACM International Conference on Human Factors in Computing Systems. Pages 2257-2266. ACM Press (2012).

Table of Figures

Table of Tables

Appendix A: Familiarity and Trust Evaluation Questionnaire

The two below questionnaires ask the first and the second group respectively to submit a rating from 1 to 5 for each other group member corresponding to familiarity and trust values. All 19 participants submitted a rating for all other subjects. In total we have collected 9 x 8 + 10 x 9 = 72 + 90 = 162 ratings (out of which 19 are "5"-self-ratings).

Appendix A.1: Group1 – Familiarity and Trust Evaluation Questionnaire

Your friends! Your music! Your party!

Give yourself a score of 5. Please use NORMALIZATION! Don't use only high and low ratings! Instead of giving absolute values, give relative scores according to your subjective valuation, i.e. evaluate answers among-themselves! E.g.: The person I trust most will receive a score of 5, the one I trust least a score of 1. Give intermediate scores to everyone else (including 1 and 5).
* Required

How much do you know your friends' musical tastes? *

	1	2	3	4	5
Participant1	○	○	○	○	○
Participant2	○	○	○	○	○
Participant3	○	○	○	○	○
Participant4	○	○	○	○	○
Participant5	○	○	○	○	○
Participant6	○	○	○	○	○
Participant7	○	○	○	○	○
Participant8	○	○	○	○	○
Participant9	○	○	○	○	○

How much do you trust your friends (music-wise)? *

	1	2	3	4	5
Participant1					
Participant2					
Participant3					
Participant4					
Participant5					
Participant6					
Participant7					
Participant8					
Participant9					

Your initials. *

Appendix A.2: Group1 – Familiarity and Trust Evaluation Questionnaire

Your friends! Your music! Your party!

Give yourself a score of 5. Please use NORMALIZATION! Don't use only high and low ratings! Instead of giving absolute values, give relative scores according to your subjective valuation, i.e. evaluate answers among-themselves! E.g.: The person I trust most will receive a score of 5, the one I trust least a score of 1. Give intermediate scores to everyone else (including 1 and 5).
* Required

How much do you know your friends' musical tastes? *

	1	2	3	4	5
Participant10					
Participant11					

	1	2	3	4	5
Participant12	○	○	○	○	○
Participant13	○	○	○	○	○
Participant14	○	○	○	○	○
Participant15	○	○	○	○	○
Participant16	○	○	○	○	○
Participant17	○	○	○	○	○
Participant18	○	○	○	○	○
Participant19	○	○	○	○	○

How much do you trust your friends (music-wise)? *

	1	2	3	4	5
Participant10	○	○	○	○	○
Participant11	○	○	○	○	○
Participant12	○	○	○	○	○
Participant13	○	○	○	○	○
Participant14	○	○	○	○	○
Participant15	○	○	○	○	○
Participant16	○	○	○	○	○
Participant17	○	○	○	○	○
Participant18	○	○	○	○	○
Participant19	○	○	○	○	○

Your initials. * ____________

Appendix B: Familiarity and Trust Ratings

In the tables below we present the results of the two online questionnaires. They contain all familiarity and trust ratings. We asked all of our participants to submit ratings of “5” for them-selves. For further analysis we removed those self-ratings as they did not contain any additional relevant information for our analysis.

Appendix B.1: Group1 – Familiarity and Trust Ratings

Familiarity	P1	P2	P3	P4	P5	P6	P7	P8	P9
P1		1	1	2	1	1	1	1	1
P2	1		5	3	4	5	5	3	4
P3	3	5		2	3	1	2	3	2
P4	3	1	1		1	1	1	1	1
P5	1	1	1	1		2	4	3	1
P6	1	1	1	1	3		3	2	3
P7	1	1	1	1	2	1		1	2
P8	1	1	1	1	1	1	1		1
P9	1	1	1	1	1	1	4	1	

Trust	P1	P2	P3	P4	P5	P6	P7	P8	P9
P1		2	2	3	2	2	2	2	2
P2	2		5	1	4	5	5	2	3
P3	3	5		3	4	1	1	3	1
P4	5	3	3		3	3	3	3	3
P5	1	1	1	1		2	4	3	2
P6	1	2	1	1	3		2	4	2
P7	1	1	1	1	3	1		1	2
P8	1	1	1	1	3	2	3		2
P9	1	1	1	1	1	1	4	1	

Appendix B.2: Group2 – Familiarity and Trust Ratings

Familiarity	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19
P10		1	1	1	2	1	1	3	3	2
P11	3		3	3	3	3	3	3	3	3
P12	3	1		2	1	4	1	3	3	1
P13	2	1	4		1	4	1	3	3	2
P14	4	2	3	3		4	3	4	2	2
P15	1	1	1	1	1		1	1	1	1
P16	3	3	3	3	3	3		3	3	3
P17	3	3	3	3	3	3	3		3	3
P18	1	1	1	1	1	1	1	1		2
P19	4	4	2	2	2	2	2	4	4	

Trust	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19
P10		1	1	1	2	1	1	3	3	2
P11	3		3	3	3	3	3	3	3	3
P12	3	1		2	1	4	1	3	3	1
P13	2	1	4		1	4	1	3	3	2
P14	4	2	3	3		4	3	4	2	2
P15	1	1	1	1	1		1	1	1	1
P16	3	3	3	3	3	3		3	3	3
P17	3	3	3	3	3	3	3		3	3
P18	1	1	1	1	1	1	1	1		2
P19	4	4	2	2	2	2	2	4	4	

Appendix C: GroupFun Contributions

In the tables below we include all the songs' names, artists and corresponding YouTube URLs contributed during the GroupFun phase of the experiment: 39 songs in Group1 and 47 songs in Group2.

Appendix C.1: Group1 – GroupFun Contributions

GroupFun order	Track	Artist	URL
1	Firework	Katy Perry	http://www.youtube.com/watch?v=QGJuMBdaqIw
2	Raise Your glass	Pink	http://www.youtube.com/watch?v=XjVNlG5cZyQ
3	Down With The Sickness	Disturbed	http://www.youtube.com/watch?v=09LTT0xwdfw
4	Engel	Rammstein	http://www.youtube.com/watch?v=I9DdyBHxGXQ
5	The Remedy	Jason Marz	http://www.youtube.com/watch?v=BW17WAwMcoQ
6	Nobody	Wonder Girls	http://www.youtube.com/watch?v=BA7fdSkp8ds
7	Someday	Flipsyde	http://www.youtube.com/watch?v=e5QEAI_O2IA
8	Lost Together	Sophie Sugar vs. Sunlounger feat. Zara	http://www.youtube.com/watch?v=lfVCpJmYdCU
9	Basket Case	Green Day	http://www.youtube.com/watch?v=NUTGr5t3MoY
10	T.N.T.	ACDC	http://www.youtube.com/watch?v=pR30knJs4Xk
11	Harvest Moon	Neil Young	http://www.youtube.com/watch?v=n2MtEsrcTTs
12	Can't Take My Eyes Off You	Frankie Valli	http://www.youtube.com/watch?v=R1j1RRWcYSg
13	System Of A Down	Chop Suey	http://www.youtube.com/watch?v=CSvFpBOe8eY
14	The Color Code	The Young Gods	http://www.youtube.com/watch?v=YoKaWCEK7dc
15	Trip In	Les Pauvres Gens	http://www.youtube.com/watch?v=AZHhUv6cdkw
16	Breaking the Habit	Linkin Park	http://www.youtube.com/watch?v=DH2KgGgK0Dc
17	Dreaming	Aloan	http://www.youtube.com/watch?v=prY4jZyw78M
18	Stirb Nicht Vor Mir	Rammstein	http://www.youtube.com/watch?v=IB8gVc4xij0
19	Long Live The Beer	The Crazy Hospital	http://www.youtube.com/watch?v=ZxCVyAwp8_w
20	Ya Basta!	Gingala	http://www.youtube.com/watch?v=rVcy_lMClA0
21	The Hand That Feeds	Nine Inch Nails	http://www.youtube.com/watch?v=xwhBRJStz7w
22	Last Moulinao	Monkey3	http://www.youtube.com/watch?v=5YrhoyvOBaU
23	Death Of Titanic	James Horner	http://www.youtube.com/watch?v=cNeC6aRuf0k
24	Deer In The Headlights	Owl City	http://www.youtube.com/watch?v=DYGNu-HBDlg
25	The Rain	Peter Luts	http://www.youtube.com/watch?v=z4VamVrF6DE
26	Heaven	Gotthard	http://www.youtube.com/watch?v=apP448jksho

27	Crave	Renaud	http://www.youtube.com/watch?v=myMDhEzDJ-g
28	Aux Champs Elysees	Joe Dassin	http://www.youtube.com/watch?v=OAMuNfs89yE
29	Aline	Herve Villard	http://www.youtube.com/watch?v=7d-Ey-0jPw8
30	Comme d'Habitude	Claude Francois	http://www.youtube.com/watch?v=tBY9jIoNzTU
31	Shera di Kaum	Ludacris	http://www.youtube.com/watch?v=eV5wfb8O8ts
32	Pulli Phirdi	Lehmber Hussainpuri	http://www.youtube.com/watch?v=FmdMru-XcNA
33	Hoor Nai Peeni Daru	Lehmber Hussainpuri	http://www.youtube.com/watch?v=W7nLrpn4csY
34	The Reward Is The Cheese	Deadmau5	http://www.youtube.com/watch?v=7QY087RDlxg
35	Sky And Sand	Paul Klankbrener	http://www.youtube.com/watch?v=1H5loYi6wVc
36	I Know You Want Me	Pitbull	http://www.youtube.com/watch?v=DnWrWSCoGis
37	Meaw	Deadmau5	http://www.youtube.com/watch?v=bn5CtUEgKMo
38	Skinny Puppy	Immortal	http://www.youtube.com/watch?v=wh1ZBIXUEPg
39	Duel Of The Fates	The London Symphony Orchestra	http://www.youtube.com/watch?v=J1gH_cjdb60

Appendix C.2: Group2 – GroupFun Contributions

The 10 members in Group2 contributed with the following songs to their group.

GroupFun order	Track	Artist	URL
1	Breaking The Habit	Linkin Park	http://www.youtube.com/watch?v=DH2KgGgK0Dc
2	The Rain	Peter Luts	http://www.youtube.com/watch?v=z4VamVrF6DE
3	Insomniac	C.I.A.	http://www.youtube.com/watch?v=cvniXGtMwsk
4	Long Live The Beer	The Crazy Hospital	http://www.youtube.com/watch?v=ZxCVyAwp8_w
5	Wake Me Up When September Ends	Green Day	http://www.youtube.com/watch?v=LqIndOEE00c
6	Paradise	Coldplay	http://www.youtube.com/watch?v=1G4isv_Fylg
7	Every Teardrop Is A Waterfall	Coldplay	http://www.youtube.com/watch?v=fyMhvkC3A84
8	Someone Like You	Adele	http://www.youtube.com/watch?v=hLQl3WQQoQ0
9	Rolling In The Deep	Adele	http://www.youtube.com/watch?v=rYEDA3JcQqw
10	Basket Case	Green Day	http://www.youtube.com/watch?v=NUTGr5t3MoY
11	Angeles	Elliott Smith	http://www.youtube.com/watch?v=FMSU4QDbdew
12	I'll Be Your Mirror	The Velvet Underground	http://www.youtube.com/watch?v=an9DoVBHHr8
13	Infra-Red	Placebo	http://www.youtube.com/watch?v=fISvc-yUU1A
14	Life Is Wonderful	Jason Mraz	http://www.youtube.com/watch?v=N7dGXBeTalY
15	Forty Feet	Franz Ferdinand	http://www.youtube.com/watch?v=wnjZXrOgvzg
16	Seize The Day	Wax Tailor	http://www.youtube.com/watch?v=TuMMKgco_c4

17	City of Blinding Lights	U2	http://www.youtube.com/watch?v=8xQOb51qZ-c
18	Paradise	Coldplay	http://www.youtube.com/watch?v=1G4isv_Fylg
19	Beautiful Day	U2	http://www.youtube.com/watch?v=co6WMzDOh1o
20	Porcelain	Moby	http://www.youtube.com/watch?v=TDrfXVwHk9Y
21	Broken Strings	James Morrison & Nelly Furtado	http://www.youtube.com/watch?v=26PAgklYYvo
22	Bright Side	Lissie	http://www.youtube.com/watch?v=1s95kupOuFE
23	This Is The Life!	Amy Macdonald	http://www.youtube.com/watch?v=iRYvuS9OxdA
24	Highway To Hell	ACDC	http://www.youtube.com/watch?v=bNlNZ2T9EeY
25	Autumn Song	Manic Street Preachers	http://www.youtube.com/watch?v=Cw9tazfA3aY
26	Adeline	Timpuri Noi	http://www.youtube.com/watch?v=p9D-RMZSFss
27	Bitter Sweet Symphony	The Verve	http://www.youtube.com/watch?v=1lyu1KKwC74
28	Na Dige In Vase Ma Del Nemishe	Hossein Bakhtiary	http://www.youtube.com/watch?v=gKN_ytldLjY
29	Khazaane Eshgh	Hossein Bakhtiary	http://www.youtube.com/watch?v=-Mi40A7EwE8
30	Baazi	Siavash Ghomayshi	http://www.youtube.com/watch?v=BPic9FYlM2I
31	Hello	Evanescence	http://www.youtube.com/watch?v=Ih61MJ72v1Y
32	Don't Worry Be Happy	Bob Marley	http://www.youtube.com/watch?v=d-diB65scQU
33	Wonderful World	James Morrison	http://www.youtube.com/watch?v=OuoaKai_L00
34	I Fell It All	Feist	http://www.youtube.com/watch?v=l-iAS18rv68
35	On My Hands	Hugh Coltman	http://www.youtube.com/watch?v=WgHcvNBp1lA
36	It Was Written	Damian Marley	http://www.youtube.com/watch?v=awYVlAv6Cek
37	Poker Face	Lady GaGa	http://www.youtube.com/watch?v=bESGLojNYSo
38	Highway To Hell Lyrics	ACDC	http://www.youtube.com/watch?v=Xv24N8H1KyI
39	Grenade	Bruno Mars	http://www.youtube.com/watch?v=SR6iYWJxHqs
40	I'll Be Your Mirror	The Velvet Underground	http://www.youtube.com/watch?v=CvoWf1HPFYM
41	It Will Rain	Bruno Mars	http://www.youtube.com/watch?v=W-w3WfgpcGg
42	Duel Of The Fates	The London Symphony Orchestra	http://www.youtube.com/watch?v=J1gH_cjdb60
43	Back To Life	Olivia Ong	http://www.youtube.com/watch?v=9KGSw_mMprk
44	Sing	Olivia Ong	http://www.youtube.com/watch?v=iVo_z0n2ONU
45	Invisible Wing	Olivia Ong	http://www.youtube.com/watch?v=5PFQfixfLoo
46	Baby	Justin Bieber	http://www.youtube.com/watch?v=kffacxfA7G4
47	Firework	Katy Perry	http://www.youtube.com/watch?v=QGJuMBdaqIw

Appendix D: Eye-tracking Study Images

These images showcase the Tobii 1750 eye-tracker screen, the iPod shuffle device, as well as the participants' position while using the eye-tracker. In the first picture (top-left) we show how we used a support extension to fixate the participants' head position during the 30 minutes duration of the experiment. Additionally, the head was positioned at the same distance from the infra-red sensor and screen. Using this extension the gaze-capture efficiency of the eye-tracker was the biggest.

Appendix E: Selection of 24 Eye-Tracking Songs

The two following tables contain the selection of 24 songs (out of the 39 and 47 GroupFun contributed songs from above) used during the eye-tracking phase of the experiment.

Appendix E.1: Group1 – Selection of 24 Eye-Tracking Songs

Order	Track	Artist
1	The Hand That Feeds	Nine Inch Nails
2	Deer In The Headlights	Owl City
3	Raise Your glass	Pink
4	Engel	Rammstein
5	Stirb Nicht Vor Mir	Rammstein
6	System Of A Down	Chop Suey
7	Lost Together	Sophie Sugar vs. Sunlounger feat. Zara
8	Nobody	Wonder Girls
9	The Color Code	The Young Gods
10	T.N.T.	ACDC
11	Dreaming	Aloan
12	Long Live The Beer	The Crazy Hospital
13	Down With The Sickness	Disturbed
14	Someday	Flipsyde
15	Can't Take My Eyes Off You	Frankie Valli
16	Ya Basta!	Gingala
17	Basket Case	Green Day
18	Death Of Titanic	James Horner
19	The Remedy	Jason Marz
20	Firework	Katy Perry
21	Trip In	Les Pauvres Gens
22	Breaking the Habit	Linkin Park
23	Last Moulinao	Monkey3
24	Harvest Moon	Neil Young

Appendix E.2: Group2 – Selection of 24 Eye-Tracking Songs

Order	Track	Artist
1	Baazi	Siavash Ghomayshi
2	Forty Feet	Franz Ferdinand
3	Basket Case	Green Day
4	Life Is Wonderful	Jason Mraz
5	Bright Side	Lissie
6	Infra-Red	Placebo
7	I'll Be Your Mirror	The Velvet Underground
8	Breaking The Habit	Linkin Park
9	Every Teardrop Is A Waterfall	Coldplay
10	Khazaane Eshgh	Hossein Bakhtiary
11	Na Dige In Vase Ma Del Nemishe	Hossein Bakhtiary
12	Paradise	Coldplay
13	Wake Me Up When September Ends	Green Day
14	Highway To Hell	ACDC
15	Someone Like You	Adele
16	This Is The Life!	Amy Macdonald
17	Insomniac	C.I.A.
18	Baby	Justin Bieber
19	Firework	Katy Perry
20	Autumn Song	Manic Street Preachers
21	Porcelain	Moby
22	Long Live The Beer	The Crazy Hospital
23	Adeline	Timpuri Noi
24	Bitter Sweet Symphony	The Verve

Appendix F: Eye-tracking interfaces ratings

In the table below we list the average and standard deviation of all group ratings in both groups separately. The average ratings in the "all other members" condition for Group1 is 3.43 and for Group2 3.5. This represents the only difference between the two groups and is due to the one extra member in Group2. For the "5 closest members" and "2 closest members" condition the average ratings displayed in the interfaces is the same: 3.65 (standard deviation 1.17) whereas for 2 closest members the average rating is 3.5 (standard deviation 1.41). In general, the first group's baseline interface ratings had an average value of 3.52 and for the second group 3.55 with same standard deviation (1.24).

Average	**8/9**	**5**	**2**	**Average**
Group1	3.43	3.65	3.5	3.52
Group2	3.5	3.65	3.5	3.55
Standard Deviation	**8/9**	**5**	**2**	**Average**
Group1	1.14	1.17	1.41	1.24
Group2	1.13	1.17	1.41	1.24

Appendix G: Group Ratings for the 24 Songs Eye-Tracking Songs

Here we include all the 24 ratings given for the corresponding eye-tracking group interface.

Appendix G.1: 24 Group1 Ratings

The Group1 members' ratings are presented the table below.

P1	P2	P3	P4	P5	P6	P7	P8	P9
3	5	3	3	4	4	2	2	2
4	2	4	4	2	2	2	4	5
5	3	4	5	2	2	1	4	4
3	5	5	3	5	4	3	1	1
3	5	3	2	4	4	3	3	3
4	5	5	4	5	3	3	1	1
5	2	1	2	1	2	4	4	2
2	1	1	3	1	2	2	2	5
2	5	2	5	5	2	3	1	1
2	4	5	1	5	4	3	2	1
4	3	2	4	5	2	4	2	4
3	4	3	4	4	3	2	3	3
2	5	3	2	5	2	2	1	2
4	1	4	5	3	2	1	2	4
5	2	4	3	3	4	5	3	5
3	4	3	2	5	3	3	2	3
5	4	5	5	4	2	5	5	5
5	2	4	1	2	3	1	3	2
4	2	3	4	3	3	2	3	5
5	1	4	5	2	3	3	4	5
2	3	3	3	5	3	1	2	4
4	5	4	5	3	4	1	2	3
3	4	2	2	5	2	3	2	2
4	3	5	2	4	4	2	3	4

Appendix G.2: 24 Group2 Ratings

The Group2 members' ratings are presented the table below.

P10	P11	P12	P13	P14	P15	P16	P17	P18	P19
3	4	1	5	5	1	5	4	2	4
2	3	3	5	5	2	3	3	3	3
5	5	5	4	4	5	2	4	5	4
4	5	3	5	4	2	3	5	4	3
3	4	5	5	3	4	2	5	4	4
2	3	5	5	4	5	1	1	4	3
3	4	3	3	2	2	2	3	1	5
3	3	3	3	3	4	2	1	5	3
5	4	5	5	5	3	3	1	3	4
2	3	2	4	2	1	5	1	1	3
2	4	2	3	4	1	5	1	1	4
5	4	4	4	4	2	5	5	5	4
5	5	4	4	3	3	3	5	5	4
3	3	4	2	3	5	3	1	5	2
5	3	5	4	4	2	3	4	5	4
5	4	5	5	3	3	4	4	4	3
5	2	5	2	2	1	4	5	1	2
5	2	3	3	2	1	1	5	1	4
4	3	4	4	2	1	2	5	4	4
5	4	5	5	5	5	3	4	4	3
5	3	5	3	5	3	2	1	4	2
1	4	5	1	4	5	4	4	1	2
1	2	4	3	1	5	4	4	1	3
4	4	5	5	5	5	4	3	5	3

Appendix H: Total Listening Times for the 24 Songs

The next two tables highlight the total eye-tracking listening time (in seconds) for all 24 selected songs.

Appendix H.1: Group1 – Total Listening Times for the 24 Songs

Interface Songs	Listening time [sec]
S1: Nine Inch Nails - The Hand That Feeds	217
S2: Owl City - Deer In The Headlights	198
S3: Pink - Raise Your Glass	241
S4: Rammstein - Engel	439
S5: Rammstein - Stirb Nicht Vor Mir	280
S6: Chop Suey - System Of A Down	316
S7: Sophie Sugar vs. Sunlounger feat. Zara - Lost Together	496
S8: Wonder Girls - Nobody	214
S9: The Young Gods - The Color Code	602
S10: ACDC -T.N.T.	214
S11: Aloan - Dreaming	242
S12: The Crazy Hospital - Long Live The Beer	293
S13: Disturbed - Down With The Sickness	288
S14: Flipsyde - Someday	433
S15: Frankie Valli - Can't Take My Eyes Off You	203
S16: Gingala - Ya Basta!	347
S17: Green Day - Basket Case	183
S18: James Horner - Death Of Titanic	541
S19: Jason Marz - The Remedy	256
S20: Katy Perry - Firework	296
S21: Les Pauvres Gens - Trip In	221
S22: Linkin Park - Breaking The Habit	530
S23: Monkey3 - Last Moulinao	381
S24: Neil Young - Harvest Moon	302

Appendix H.2: Group2 – Total Listening Times for the 24 Songs

Interface Songs	**Watching time [sec]**
S1: Baazi Siavash - Ghomayshi	280
S2: Franz Ferdinand - Forty Feet	236
S3: Green Day - Basket Case	183
S4: Jason Mraz - Life Is Wonderful	160
S5: Lissie - Bright Side	195
S6: Placebo - Infra-Red	192
S7: The Velvet Underground - I'll Be Your Mirror	106
S8: Linkin Park - Breaking The Habit	528
S9: Coldplay - Every Teardrop Is A Waterfall	249
S10: Hossein Bakhtiary - Khazaane Eshgh	494
S11: Hossein Bakhtiary - Na Dige In Vase Ma Del Nemishe	360
S12: Coldplay - Paradise	338
S13: Green Day - Wake Me Up When September Ends	283
S14: ACDC - Highway To Hell	291
S15: Adele - Someone Like You	311
S16: Amy Macdonald - This Is The Life!	293
S17: C.I.A. - Insomniac	190
S18: Justin Bieber - Baby	225
S19: Katy Perry - Firework	234
S20: Manic Street Preachers - Autumn Song	310
S21: Moby - Porcelain	229
S22: The Crazy Hospital - Long Live The Beer	293
S23: Timpuri Noi - Adeline	233
S24: The Verve - Bitter Sweet Symphony	381

Appendix I: Visualization Statistics Summary

In this section we present a summary of visualization statistics per group. Visualization times in each area of interest (AOIS: AOI Songs, AOIU: AOI Users and AOIR: AOI Ratings) are displayed with the following abbreviations: TTI = Total Time Indicated by the eye-tracker and TTC = Total Time Computed (in the interface).

Appendix I.1: Group1 – Visualization Statistics Summary

Group1	TTI	TTC	Time in AOIS	Time in AOIU	Time in AOIR	< 5 sec	< 10 sec
Seconds	4867.5	4080.7	1536.7	1382.2	1161.8	37	95
Minutes	81.13	68.01	25.61	23.04	19.36	456	456
Hours	1.352	1.134	0.427	0.384	0.323		
Percentage		100%	38%	34%	28%	8.11%	20.83%

The table above contains the sum of total time indicated (second column) and the total time computed (third column) split by "Time in AOI Songs", "Time in AOI Users" and "Time in AOI Ratings". We converted TTI and TTC from seconds into minutes and hours in order to highlight the large amount of data collection time. Besides this representation, all eye-tracking information account for more than 2GB of output files!

In Group1 the total time indicated is around 81 minutes continuous recording. Out of this value, around 68 minutes were fully captured through the relevant fixation times in the eye-gaze interfaces. In summary, 38% of the total time (about 25 minutes) subjects spend by looking at AOI Songs, 34% (about 23 minutes) at AOI Users and 28% (about 19 minutes) at AOI Ratings.

The last 2 columns show the number of interfaces (out of the total number of 456) for which participants spent less than 5 seconds (37 out of 456, i.e. 8.11%) and less than 10 seconds (95 out of 456, i.e. 20.83%). This result proves that indeed participants focused their attention on the displayed information analyzing the song information, participants' names and their ratings shown to them.

Appendix I.2: Group2 – Visualization Statistics Summary

Group2	TTI	TTC	Time in AOIS	Time in AOIU	Time in AOIR	< 5 sec	< 10 sec
Seconds	6,437	5336.4	1,517	2188.9	1,630	16	81
Minutes	107.28	88.94	25.29	36.48	27.17	456	456
Hours	1.788	1.482	0.422	0.608	0.453		
Percentage		100%	28%	41%	31%	3.50%	17.76%

The total time indicated on the eye tracker for Group2 is about 107 minutes. Out of this, about 89 minutes were fully recorded using the fixation points.

In Group2 28% of the total time (about 25 minutes) subjects spend by looking at AOI Songs, 41% (about 36 minutes) at AOI Users and 31% (about 27 minutes) at AOI Ratings.

The last 2 columns show the number of interfaces (out of the total number of 456) for which participants spent less than 5 seconds (16 out of 456, i.e. 3.5%) and less than 10 seconds (81 out of 456, i.e. 17.75%).

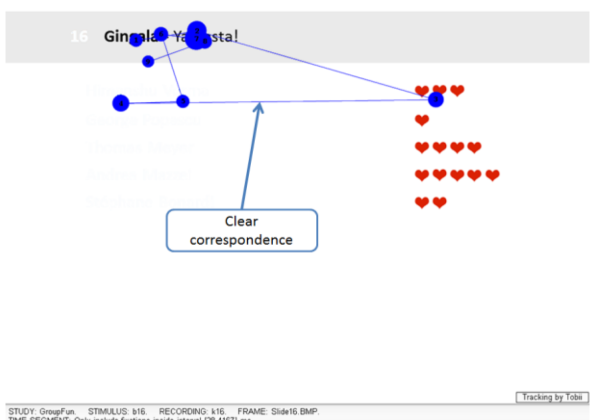
In contrast with Group1, members of Group2 spent more time by looking at the interfaces, especially for the time in AOI Users and time in AOI Ratings, thus generating more relevant horizontal correspondences. In addition to this, for only 16 interfaces (accounting for 3.5%) they gazed for less than 5 seconds and for only 81 interfaces (17.76%) they looked for less than 10 seconds. This difference is not only due to the one extra member, which does not justify the full difference, but also for the members' interest to see others' preferences when inferring a self-decision.

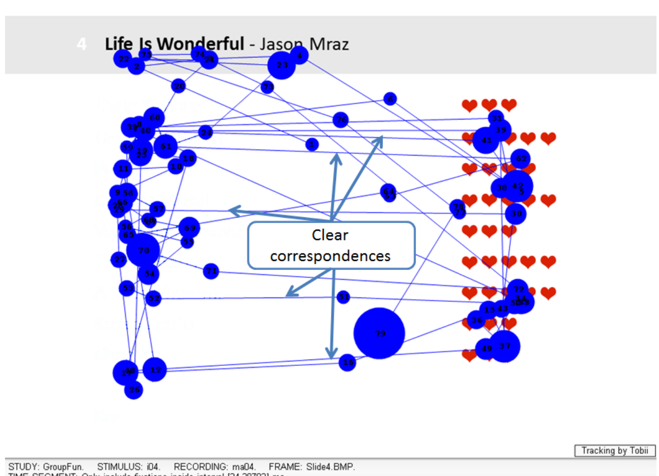
Appendix J: Eye-Gaze

In the following we present examples of clear, unclear and complex eye-gaze.

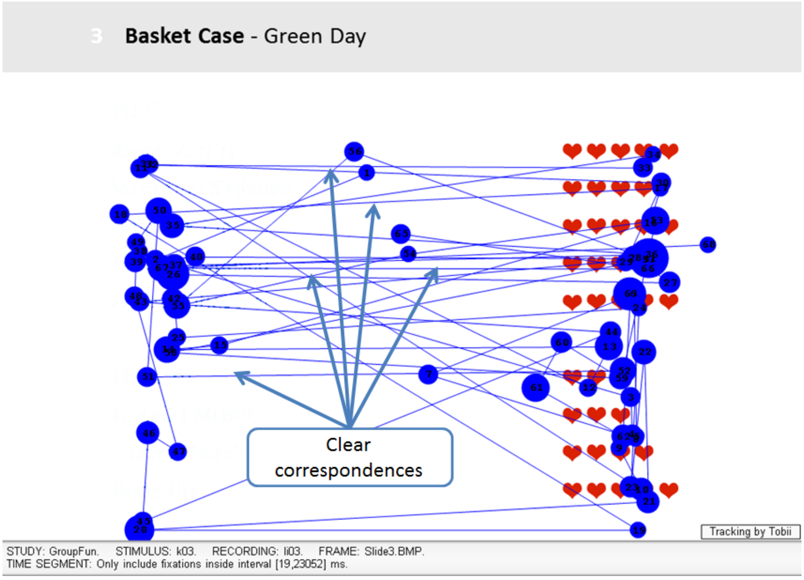
Appendix J.1: Clear Eye-Gaze

In the first figure below clear horizontal correspondence between the name of one participant (removed for privacy reasons) to the left and his rating to the right can be observed. In the second one, the participant observed most of the information displayed to him and 5 clear correspondences can be noticed.

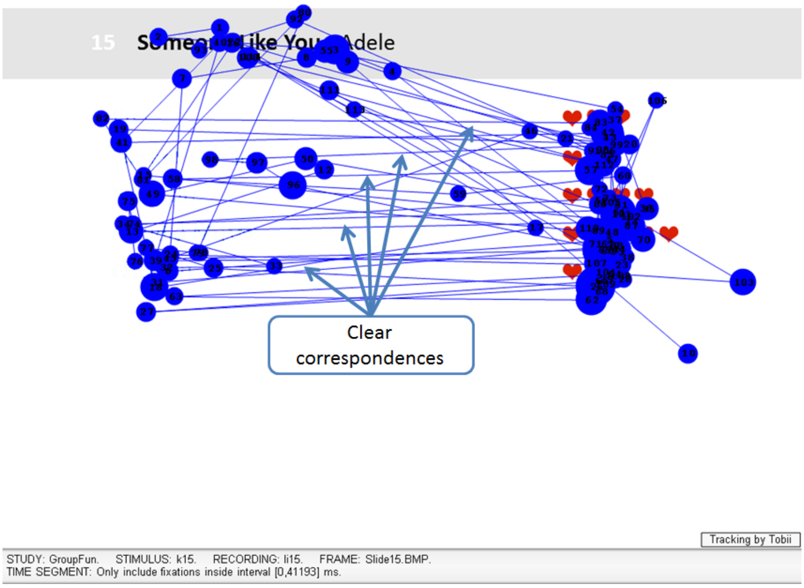




In this screenshot the participant created a big number of horizontal correspondences between the names of his peers and their ratings. These can be clearly seen.

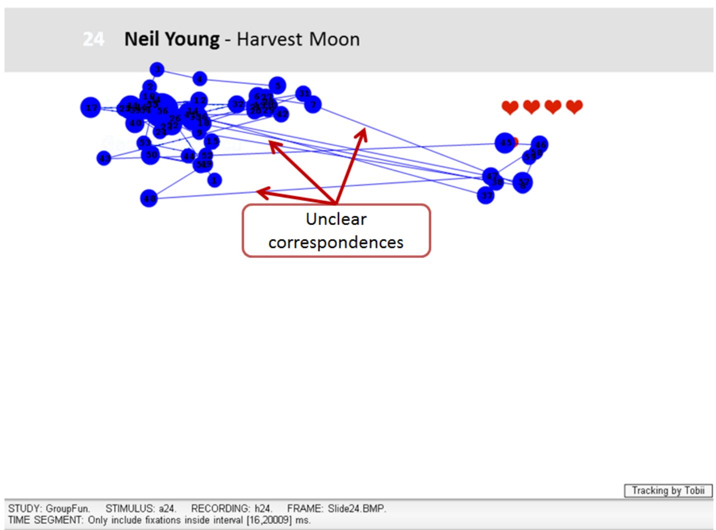


In the below interface displaying 5 closest members, the current participant created all possible horizontal correspondences, thus being influenced by all.

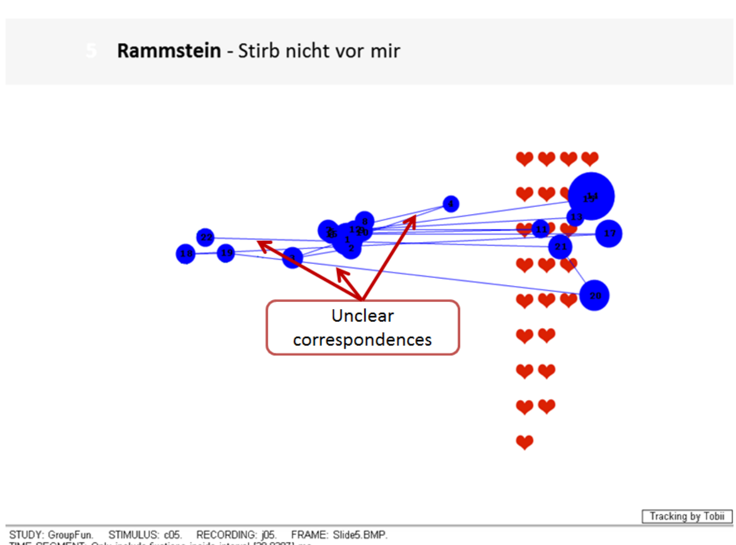


Appendix J.2: Unclear Eye-Gaze

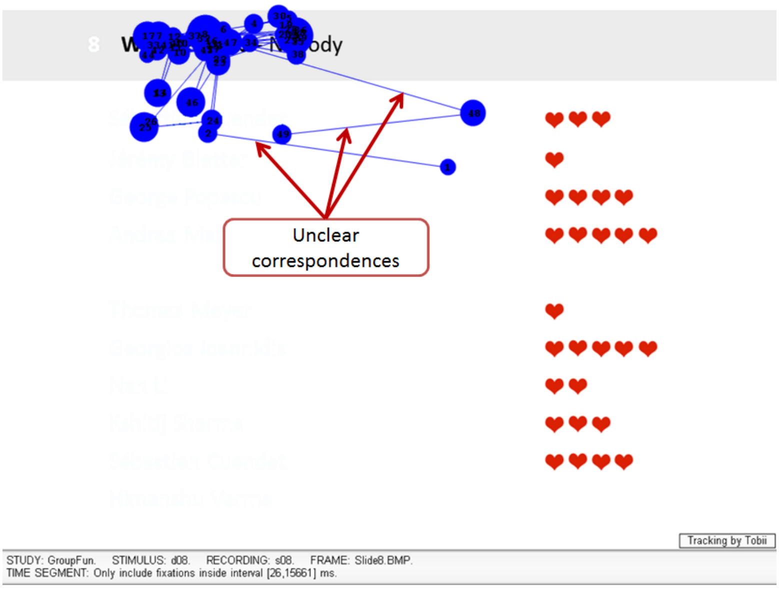
In this interface presenting 2 closest members, few unclear correspondences can be seen: lines are skewed from the horizontal axis.



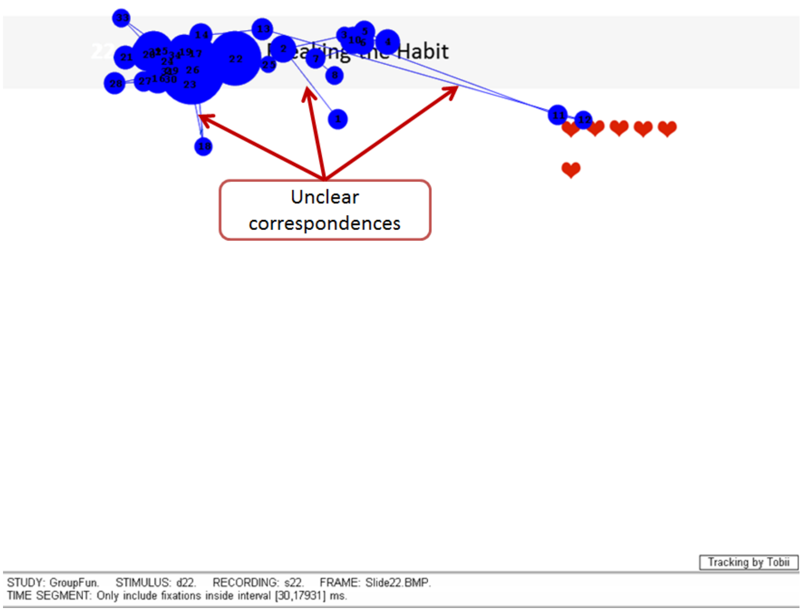
In the screenshot below the focus of the viewer is set in the middle of the screen with no particular connection between other users and ratings.



The current user looked only at the information displayed to the top-right of the screen which mostly includes the song's name and one peer name.

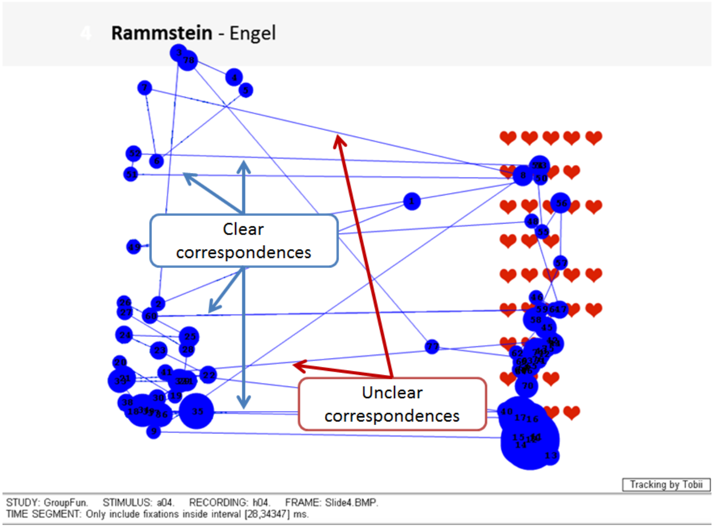


As before, the unclear correspondences presented in the figure indicate an eye-gaze focused on the song's name rather than on users and ratings.

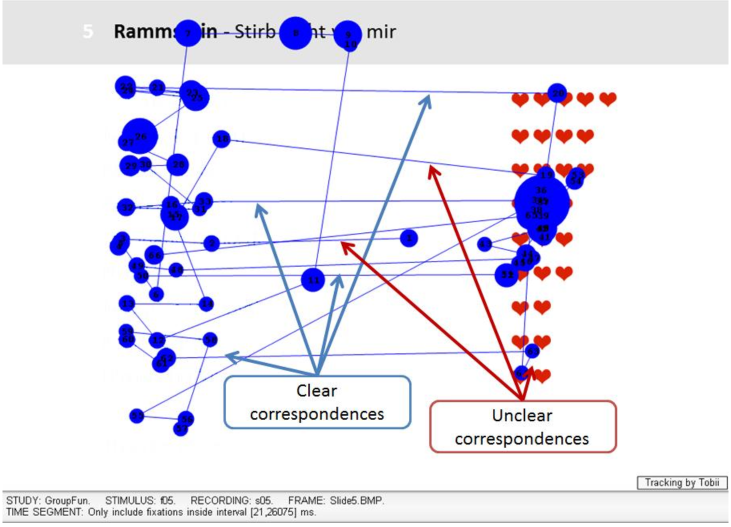


Appendix J.3: Complex Eye-Gaze

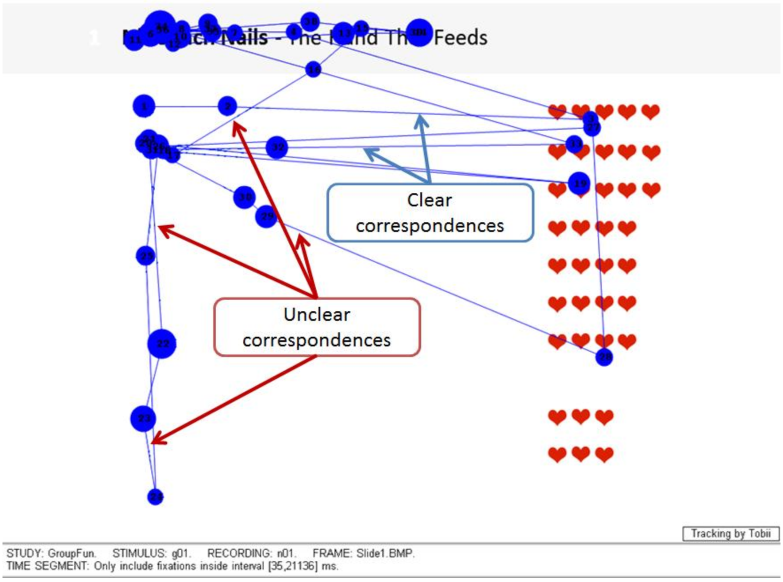
Both clear and unclear correspondences are recorded in the eye-tracker's output in the figure above.



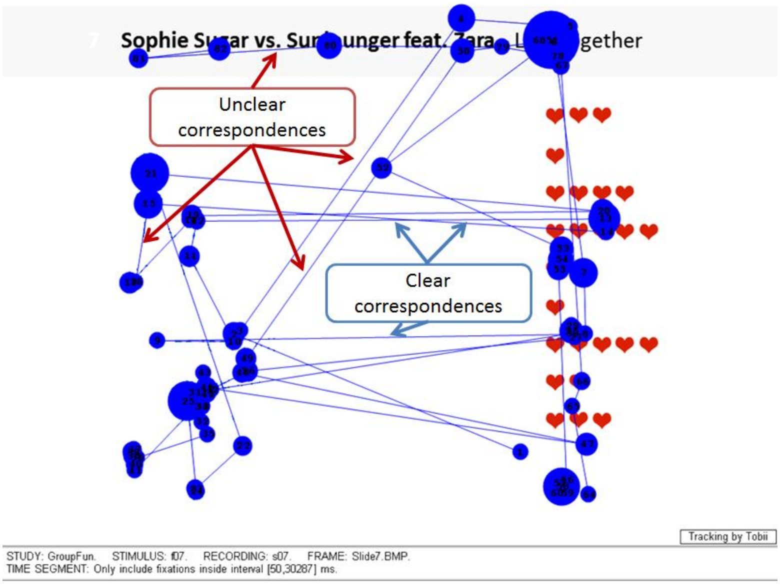
Despite the many correspondences that the current participant produces, there are only 4 horizontal ones and other few skewed lines.



Both types of lines can take place in the same gaze: the unclear ones connect users' names whereas the almost horizontal ones users with their ratings.



The eye-gaze lines between regions of the same AOI (e.g. AOIS) are unclear. Clear gaze is presented in multiple parallel lines.



Appendix K: Top3 participants

In this section we include the results from extracting the Top3 closest participants for each subject based on several criteria: music preference or 24 ratings correlation, familiarity, trust and eye-tracking data. For some parameters, such as the ratings correlation and eye-gaze information all Top3 closest participants could be determined. For the other two parameters, familiarity and trust, the empty spaces in the tables are due to the fact that subjects did not differentiate for first level, second level, or third level familiarity or trust relationships.

Appendix K.1: Group1 – Top3 participants

Top 3 based on 24 ratings correlation			
P1	P2	P4	P7
P2	P1	P4	P7
P3	P8	P6	P7
P4	P7	P2	P1
P5	P3	P8	P1
P6	P8	P3	P9
P7	P4	P3	P2
P8	P3	P6	P5
P9	P6	P3	P8

In the table above P5's music preference is most alike with P3, P8 and P1. However, from the table below we report that he is most familiar with P7, P8 and P6's music preferences and trusts P7, P8 and P6 (in the same order). Moreover, the eye-gaze shows that P5 looked most at P7, P2 and P8's ratings in the 24 interfaces.

Top 3 based on Familiarity			
P1	P4		
P2	P3	P6	P7
P3	P2	P5	P1
P4	P1		
P5	P7	P8	P6
P6	P5	P7	P9
P7	P5	P9	
P8			
P9	P7		

Top 3 based on Trust			
P1	P4		
P2	P3	P6	P7
P3	P2	P5	P1
P4	P1		
P5	P7	P8	P6
P6	P5	P7	P9
P7	P5	P9	
P8	P5	P7	P9
P9	P7		

Top 3 based on Eye-Tracking			
P1	P4	P9	P2
P2	P5	P4	P7
P3	P9	P2	P1
P4	P7	P5	P6
P5	P7	P2	P8
P6	P5	P7	P8
P7	P5	P4	P2
P8	P5	P2	P7
P9	P7	P5	P2

Appendix K.2: Group2 – Top3 participants

Now, we take as example P18: from the tables below we report that this participant misses familiarity and trust values for second and third most familiar and trusted members. Indeed, from the familiarity and trust rating tables we report that P18 submitted only one rating of "2" for both familiarity and trust for P19 and all other ratings of "1" for all other members. For this subject, the 24 ratings show him to be closest to P10, P12 and P13 (in this order) and the eye-gaze show him to look most at P19, P17 and P14.

Top 3 based on 24 ratings correlation			
P10	P18	P12	P17
P11	P14	P13	P18
P12	P15	P18	P10
P13	P14	P18	P11
P14	P11	P13	P18
P15	P12	P18	P14
P16	P11	P14	P17
P17	P10	P19	P11
P18	P10	P12	P13
P19	P13	P11	P17

Top 3 based on Familiarity			
P10	P17	P18	P14
P11			
P12	P15	P10	P17
P13	P15	P12	P17
P14	P10	P15	P17
P15			
P16			
P17			
P18	P19		
P19	P17	P18	P11

Top 3 based on Trust			
P10	P17	P18	P14
P11			
P12	P15	P10	P17
P13	P15	P12	P17
P14	P10	P12	P17
P15			
P16			
P17			
P18	P19		
P19	P17	P18	P11

Top 3 based on Eye-Tracking			
P10	P14	P17	P13
P11	P17	P19	P15
P12	P15	P13	P14
P13	P15	P12	P11
P14	P10	P15	P13
P15	P10	P12	P14
P16	P15	P14	P13
P17	P14	P10	P19
P18	P19	P17	P14
P19	P17	P18	P11

Appendix L: Correlation Values

In this section we compute various correlation values per group. First we normalize familiarity and trust values per participant. Then, we compute an average between familiarity and trust and normalize again the new values. Next, we extract the number of total horizontal correspondences for each participant for the 24 seen interfaces and normalize into an eye-tracking score. Finally, we compute the correlation between: (1) normalized familiarity and normalized eye-gaze, (2) normalized trust and normalized eye-gaze, (3) normalized familiarity and trust and normalized eye-gaze and find out that the last one outputs highest average correlation values. Thus, we show that subjective values are highly correlated with objective eye-tracking data.

Appendix L.1: Group1 - Correlation Values

Appendix L.1.1: Familiarity and Eye-Gaze

Normalized Familiarity by Sum										
Users	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**Total**
P1		0.111111	0.111	0.22222	0.11111	0.1111	0.1111	0.1111	0.1111	1
P2	0.033333		0.167	0.1	0.13333	0.1667	0.1667	0.1	0.1333	1
P3	0.142857	0.238095		0.09524	0.14286	0.0476	0.0952	0.1429	0.0952	1
P4	0.3	0.1	0.1		0.1	0.1	0.1	0.1	0.1	1
P5	0.071429	0.071429	0.071	0.07143		0.1429	0.2857	0.2143	0.0714	1
P6	0.066667	0.066667	0.067	0.06667	0.2		0.2	0.1333	0.2	1
P7	0.1	0.1	0.1	0.1	0.2	0.1		0.1	0.2	1
P8	0.125	0.125	0.125	0.125	0.125	0.125	0.125		0.125	1
P9	0.090909	0.090909	0.091	0.09091	0.09091	0.0909	0.3636	0.0909		1

Normalized Familiarity by Sum and Normalized Eye-Gaze / Average = 0.510, Std. = 0.310										
Users	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**Corr.**
P1		0.111111	0.111	0.22222	0.11111	0.1111	0.1111	0.1111	0.1111	0.72672
P1		0.153846	0	0.33333	0.07692	0	0.1282	0.0769	0.2308	
P2	0.033333		0.167	0.1	0.13333	0.1667	0.1667	0.1	0.1333	0.27735
P2	0.125		0.1	0.175	0.275	0.05	0.125	0.075	0.075	
P3	0.142857	0.238095		0.09524	0.14286	0.0476	0.0952	0.1429	0.0952	0.5944
P3	0.176471	0.176471		0.17647	0.11765	0	0	0.1176	0.2353	
P4	0.3	0.1	0.1		0.1	0.1	0.1	0.1	0.1	-0.1138
P4	0	0	0		0.28571	0.1429	0.4286	0	0.1429	
P5	0.071429	0.071429	0.071	0.07143		0.1429	0.2857	0.2143	0.0714	0.63489
P5	0.108108	0.189189	0.108	0.10811		0.027	0.2162	0.1622	0.0811	
P6	0.066667	0.066667	0.067	0.06667	0.2		0.2	0.1333	0.2	0.8455
P6	0.085106	0.12766	0	0.06383	0.25532		0.2128	0.1277	0.1277	
P7	0.1	0.1	0.1	0.1	0.2	0.1		0.1	0.2	0.34597
P7	0.08	0.24	0	0.24	0.24	0.08		0.08	0.04	
P8	0.125	0.125	0.125	0.125	0.125	0.125	0.125		0.125	0.42776
P8	0.067797	0.186441	0	0.11864	0.30508	0.0678	0.1695		0.0847	
P9	0.090909	0.090909	0.091	0.09091	0.09091	0.0909	0.3636	0.0909		0.84688
P9	0.083333	0.145833	0	0.10417	0.1875	0.0833	0.3125	0.0833		

Appendix L.1.2: Trust and Eye-Gaze

Normalized Trust by Sum										
Users	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**Total**
P1		0.117647	0.118	0.17647	0.11765	0.1176	0.1176	0.1176	0.1176	1
P2	0.074074		0.185	0.03704	0.14815	0.1852	0.1852	0.0741	0.1111	1
P3	0.142857	0.238095		0.14286	0.19048	0.0476	0.0476	0.1429	0.0476	1
P4	0.192308	0.115385	0.115		0.11538	0.1154	0.1154	0.1154	0.1154	1
P5	0.066667	0.066667	0.067	0.06667		0.1333	0.2667	0.2	0.1333	1
P6	0.0625	0.125	0.063	0.0625	0.1875		0.125	0.25	0.125	1
P7	0.090909	0.090909	0.091	0.09091	0.27273	0.0909		0.0909	0.1818	1
P8	0.071429	0.071429	0.071	0.07143	0.21429	0.1429	0.2143		0.1429	1
P9	0.090909	0.090909	0.091	0.09091	0.09091	0.0909	0.3636	0.0909		1

Correlation between Normalized Trust and Normalized Eye-Gazes from A to B? Avg.=0.529, Std.=0.245										
Users	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**Corr.**
P1		0.117647	0.118	0.17647	0.11765	0.1176	0.1176	0.1176	0.1176	0.61975
P1		0.153846	0	0.33333	0.07692	0	0.1282	0.0769	0.2308	
P2	0.074074		0.185	0.03704	0.14815	0.1852	0.1852	0.0741	0.1111	0.2321
P2	0.125		0.1	0.175	0.275	0.05	0.125	0.075	0.075	
P3	0.142857	0.238095		0.14286	0.19048	0.0476	0.0476	0.1429	0.0476	0.55596
P3	0.176471	0.176471		0.17647	0.11765	0	0	0.1176	0.2353	
P4	0.192308	0.115385	0.115		0.11538	0.1154	0.1154	0.1154	0.1154	0.07008
P4	0	0	0		0.28571	0.1429	0.4286	0	0.1429	
P5	0.066667	0.066667	0.067	0.06667		0.1333	0.2667	0.2	0.1333	0.59745
P5	0.108108	0.189189	0.108	0.10811		0.027	0.2162	0.1622	0.0811	
P6	0.0625	0.125	0.063	0.0625	0.1875		0.125	0.25	0.125	0.70389
P6	0.085106	0.12766	0	0.06383	0.25532		0.2128	0.1277	0.1277	
P7	0.090909	0.090909	0.091	0.09091	0.27273	0.0909		0.0909	0.1818	0.43988
P7	0.08	0.24	0	0.24	0.24	0.08		0.08	0.04	
P8	0.071429	0.071429	0.071	0.07143	0.21429	0.1429	0.2143		0.1429	0.69132
P8	0.067797	0.186441	0	0.11864	0.30508	0.0678	0.1695		0.0847	
P9	0.090909	0.090909	0.091	0.09091	0.09091	0.0909	0.3636	0.0909		0.84688
P9	0.083333	0.145833	0	0.10417	0.1875	0.0833	0.3125	0.0833		

Appendix L.1.3: Familiarity, Trust and Eye-Gaze

Average Familiarity and Trust

Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9	Total
P1		1.5	1.5	2.5	1.5	1.5	1.5	1.5	1.5	13
P2	1.5		5	2	4	5	5	2.5	3.5	28.5
P3	3	5		2.5	3.5	1	1.5	3	1.5	21
P4	4	2	2		2	2	2	2	2	18
P5	1	1	1	1		2	4	3	1.5	14.5
P6	1	1.5	1	1	3		2.5	3	2.5	15.5
P7	1	1	1	1	2.5	1		1	2	10.5
P8	1	1	1	1	2	1.5	2		1.5	11
P9	1	1	1	1	1	1	4	1		11

Normalized familiarity and trust = social distance

Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9	Total
P1		0.115385	0.115	0.19231	0.11538	0.1154	0.1154	0.1154	0.1154	1
P2	0.052632		0.175	0.07018	0.14035	0.1754	0.1754	0.0877	0.1228	1
P3	0.142857	0.238095		0.11905	0.16667	0.0476	0.0714	0.1429	0.0714	1
P4	0.222222	0.111111	0.111		0.11111	0.1111	0.1111	0.1111	0.1111	1
P5	0.068966	0.068966	0.069	0.06897		0.1379	0.2759	0.2069	0.1034	1
P6	0.064516	0.096774	0.065	0.06452	0.19355		0.1613	0.1935	0.1613	1
P7	0.095238	0.095238	0.095	0.09524	0.2381	0.0952		0.0952	0.1905	1
P8	0.090909	0.090909	0.091	0.09091	0.18182	0.1364	0.1818		0.1364	1
P9	0.090909	0.090909	0.091	0.09091	0.09091	0.0909	0.3636	0.0909		1

Eye-Tracking social distance / numbers of horizontal correspondences

Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9	Total
P1		6	0	13	3	0	5	3	9	39
P2	5		4	7	11	2	5	3	3	40
P3	3	3		3	2	0	0	2	4	17
P4	0	0	0		2	1	3	0	1	7
P5	4	7	4	4		1	8	6	3	37
P6	4	6	0	3	12		10	6	6	47
P7	4	12	0	12	12	4		4	2	50
P8	4	11	0	7	18	4	10		5	59
P9	4	7	0	5	9	4	15	4		48

Normalized eye-tracking social distance										
Participant	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**Total**
P1		0.153846	0	0.33333	0.07692	0	0.1282	0.0769	0.2308	1
P2	0.125		0.1	0.175	0.275	0.05	0.125	0.075	0.075	1
P3	0.176471	0.176471		0.17647	0.11765	0	0	0.1176	0.2353	1
P4	0	0	0		0.28571	0.1429	0.4286	0	0.1429	1
P5	0.108108	0.189189	0.108	0.10811		0.027	0.2162	0.1622	0.0811	1
P6	0.085106	0.12766	0	0.06383	0.25532		0.2128	0.1277	0.1277	1
P7	0.08	0.24	0	0.24	0.24	0.08		0.08	0.04	1
P8	0.067797	0.186441	0	0.11864	0.30508	0.0678	0.1695		0.0847	1
P9	0.083333	0.145833	0	0.10417	0.1875	0.0833	0.3125	0.0833		1

Correlation between social distance between A and B and gazes from A to B? Average = 0.546, Std.=0.278										
Participant	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**Corr.**
P1		0.115	0.115	0.192	0.115	0.115	0.115	0.115	0.115	0.668
P1		0.154	0	0.333	0.077	0	0.128	0.077	0.231	
P2	0.053		0.175	0.07	0.14	0.175	0.175	0.088	0.123	0.261
P2	0.125		0.1	0.175	0.275	0.05	0.125	0.075	0.075	
P3	0.143	0.238		0.119	0.167	0.048	0.071	0.143	0.071	0.587
P3	0.176	0.176		0.176	0.118	0	0	0.118	0.235	
P4	0.222	0.111	0.111		0.111	0.111	0.111	0.111	0.111	0
P4	0	0	0		0.286	0.143	0.429	0	0.143	
P5	0.069	0.069	0.069	0.069		0.138	0.276	0.207	0.103	0.622
P5	0.108	0.189	0.108	0.108		0.027	0.216	0.162	0.081	
P6	0.065	0.097	0.065	0.065	0.194		0.161	0.194	0.161	0.842
P6	0.085	0.128	0	0.064	0.255		0.213	0.128	0.128	
P7	0.095	0.095	0.095	0.095	0.238	0.095		0.095	0.19	0.407
P7	0.08	0.24	0	0.24	0.24	0.08		0.08	0.04	
P8	0.091	0.091	0.091	0.091	0.182	0.136	0.182		0.136	0.684
P8	0.068	0.186	0	0.119	0.305	0.068	0.169		0.085	
P9	0.091	0.091	0.091	0.091	0.091	0.091	0.364	0.091		0.847
P9	0.083	0.146	0	0.104	0.188	0.083	0.313	0.083		

Appendix L.2: Group2 - Correlation Values

Appendix L.2.1: Familiarity and Eye-Gaze

Normalized Familiarity by Sum											
Participant	**P10**	**P11**	**P12**	**P13**	**P14**	**P15**	**P16**	**P17**	**P18**	**P19**	**Total**
P10		0.083	0.083	0.083	0.0833	0.083	0.0833	0.333	0.083	0.083	1
P11	0.176		0.058	0.058	0.0588	0.117	0.0588	0.176	0.117	0.176	1
P12	0.076	0.076		0.153	0.0769	0.230	0.076	0.076	0.153	0.076	1
P13	0.090	0.090	0.272		0.0909	0.090	0.0909	0.090	0.090	0.090	1
P14	0.190	0.095	0.095	0.095		0.095	0.1428	0.047	0.095	0.142	1
P15	0.111	0.111	0.111	0.111	0.1111		0.1111	0.111	0.111	0.111	1
P16	0.111	0.111	0.111	0.111	0.1111	0.111		0.111	0.111	0.111	1
P17	0.111	0.111	0.111	0.111	0.1111	0.111	0.1111		0.111	0.111	1
P18	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		0.2	1
P19	0.133	0.133	0.066	0.066	0.0666	0.066	0.0666	0.133	0.266		1

Normalized Familiarity by Sum and Normalized Eye-Gaze / Avg.=0.581, Std.=0.241											
Participant	**P10**	**P11**	**P12**	**P13**	**P14**	**P15**	**P16**	**P17**	**P18**	**P19**	**Corr.**
P10		0.083	0.083	0.083	0.0833	0.083	0.083	0.333	0.083	0.083	0.591
P10		0.081	0.114	0.147	0.196	0.098	0.065	0.180	0.081	0.032	
P11	0.176		0.058	0.058	0.058	0.117	0.0588	0.176	0.117	0.176	0.874
P11	0.115		0.073	0.063	0.031	0.147	0.063	0.242	0.105	0.157	
P12	0.076	0.076		0.153	0.076	0.230	0.0769	0.076	0.153	0.076	0.744
P12	0.154	0.056		0.169	0.169	0.239	0.070	0.028	0.084	0.028	
P13	0.090	0.090	0.272		0.090	0.090	0.090	0.090	0.090	0.090	0.166
P13	0	0.111	0.111		0.111	0.333	0	0.111	0	0.222	
P14	0.190	0.095	0.095	0.095		0.095	0.142	0.047	0.095	0.142	0.501
P14	0.236	0.052	0.078	0.157		0.157	0.052	0.131	0.078	0.052	
P15	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	0.111	0.313
P15	0.307	0	0.230	0.153	0.192		0.038	0.038	0	0.038	
P16	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	0.394
P16	0.122	0.052	0.052	0.175	0.210	0.263		0.070	0.035	0.017	
P17	0.111	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.564
P17	0.169	0.094	0.056	0.056	0.207	0.093	0.056		0.113	0.150	
P18	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		0.2	0.802
P18	0.064	0.129	0	0.129	0.129	0.032	0.096	0.129		0.290	
P19	0.133	0.133	0.066	0.066	0.066	0.066	0.066	0.133	0.266		0.863
P19	0.135	0.162	0.054	0.054	0.081	0.054	0.040	0.216	0.202		

Appendix L.2.2: Trust and Eye-Gaze

Normalized Trust by Sum

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Total
P10		0.067	0.067	0.067	0.133	0.067	0.067	0.2	0.2	0.133	1
P11	0.111		0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	1
P12	0.158	0.053		0.105	0.053	0.211	0.053	0.158	0.158	0.053	1
P13	0.095	0.048	0.19		0.048	0.19	0.048	0.143	0.143	0.095	1
P14	0.148	0.074	0.111	0.111		0.148	0.111	0.148	0.074	0.074	1
P15	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	0.111	1
P16	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	1
P17	0.111	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	1
P18	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		0.2	1
P19	0.154	0.154	0.077	0.077	0.077	0.077	0.077	0.154	0.154		1

Normalized Trust and Normalized Eye-Gazes from A to B / Avg.=0.585, Std.=0.198

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Corr.
P10		0.066	0.066	0.066	0.133	0.066	0.066	0.2	0.2	0.133	0.468
P10		0.081	0.114	0.147	0.196	0.098	0.065	0.180	0.081	0.032	
P11	0.111		0.111	0.111	0.111	0.111	0.1111	0.111	0.111	0.111	0.501
P11	0.115		0.073	0.063	0.031	0.147	0.063	0.242	0.105	0.157	
P12	0.157	0.052		0.105	0.052	0.210	0.052	0.157	0.157	0.052	0.575
P12	0.154	0.056		0.169	0.169	0.239	0.070	0.028	0.084	0.028	
P13	0.095	0.047	0.190		0.047	0.190	0.047	0.142	0.142	0.095	0.497
P13	0	0.111	0.111		0.111	0.333	0	0.111	0	0.222	
P14	0.148	0.074	0.111	0.111		0.148	0.111	0.148	0.074	0.074	0.814
P14	0.236	0.052	0.078	0.157		0.157	0.052	0.131	0.078	0.052	
P15	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	0.111	0.313
P15	0.307	0	0.230	0.153	0.192		0.038	0.038	0	0.038	
P16	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	0.394
P16	0.122	0.052	0.052	0.175	0.210	0.263		0.070	0.035	0.017	
P17	0.111	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.564
P17	0.169	0.094	0.056	0.056	0.207	0.09	0.056		0.113	0.150	
P18	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		0.2	0.802
P18	0.064	0.129	0	0.129	0.129	0.032	0.096	0.129		0.290	
P19	0.153	0.153	0.076	0.076	0.076	0.076	0.076	0.153	0.153		0.925
P19	0.135	0.162	0.054	0.054	0.081	0.054	0.040	0.216	0.202		

Appendix L.2.3: Familiarity, Trust and Eye-Gaze

Average Familiarity and Trust

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Total
P10		1	1	1	1.5	1	1	3.5	2	1.5	13.5
P11	3		2	2	2	2.5	2	3	2.5	3	22
P12	2	1		2	1	3.5	1	2	2.5	1	16
P13	1.5	1	3.5		1	2.5	1	2	2	1.5	16
P14	4	2	2.5	2.5		3	3	2.5	2	2.5	24
P15	1	1	1	1	1		1	1	1	1	9
P16	2	2	2	2	2	2		2	2	2	18
P17	3	3	3	3	3	3	3		3	3	27
P18	1	1	1	1	1	1	1	1		2	10
P19	3	3	1.5	1.5	1.5	1.5	1.5	3	4		20.5

Normalized familiarity and trust = social distance

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Total
P10		0.074	0.074	0.074	0.111	0.074	0.074	0.259	0.148	0.111	1
P11	0.136		0.091	0.091	0.091	0.114	0.091	0.136	0.114	0.136	1
P12	0.125	0.063		0.125	0.063	0.219	0.063	0.125	0.156	0.063	1
P13	0.094	0.063	0.219		0.063	0.156	0.063	0.125	0.125	0.094	1
P14	0.167	0.083	0.104	0.104		0.125	0.125	0.104	0.083	0.104	1
P15	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	0.111	1
P16	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	1
P17	0.111	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	1
P18	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		0.2	1
P19	0.146	0.146	0.073	0.073	0.073	0.073	0.073	0.146	0.195		1

Eye-Tracking social distance / numbers of horizontal correspondences

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Total
P10		5	7	9	12	6	4	11	5	2	61
P11	11		7	6	3	14	6	23	10	15	95
P12	11	4		12	12	17	5	2	6	2	71
P13	0	1	1		1	3	0	1	0	2	9
P14	9	2	3	6		6	2	5	3	2	38
P15	8	0	6	4	5		1	1	0	1	26
P16	7	3	3	10	12	15		4	2	1	57
P17	9	5	3	3	11	5	3		6	8	53
P18	2	4	0	4	4	1	3	4		9	31
P19	10	12	4	4	6	4	3	16	15		74

Normalized eye-tracking social distance

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Total
P10		0.082	0.115	0.148	0.197	0.098	0.066	0.18	0.082	0.033	1
P11	0.116		0.074	0.063	0.032	0.147	0.063	0.242	0.105	0.158	1
P12	0.155	0.056		0.169	0.169	0.239	0.07	0.028	0.085	0.028	1
P13	0	0.111	0.111		0.111	0.333	0	0.111	0	0.222	1
P14	0.237	0.053	0.079	0.158		0.158	0.053	0.132	0.079	0.053	1
P15	0.308	0	0.231	0.154	0.192		0.038	0.038	0	0.038	1
P16	0.123	0.053	0.053	0.175	0.211	0.263		0.07	0.035	0.018	1
P17	0.17	0.094	0.057	0.057	0.208	0.094	0.057		0.113	0.151	1
P18	0.065	0.129	0	0.129	0.129	0.032	0.097	0.129		0.29	1
P19	0.135	0.162	0.054	0.054	0.081	0.054	0.041	0.216	0.203		1

Correlation between social distance between A and B and gazes from A to B / Average = 0.625, Std.=0.205

Participant	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	Corr.
P10		0.074	0.074	0.074	0.111	0.074	0.074	0.259	0.148	0.111	0.583
P10		0.082	0.115	0.148	0.197	0.098	0.066	0.18	0.082	0.033	
P11	0.136		0.091	0.091	0.091	0.114	0.091	0.136	0.114	0.136	0.79
P11	0.116		0.074	0.063	0.032	0.147	0.063	0.242	0.105	0.158	
P12	0.125	0.063		0.125	0.063	0.219	0.063	0.125	0.156	0.063	0.68
P12	0.155	0.056		0.169	0.169	0.239	0.07	0.028	0.085	0.028	
P13	0.094	0.063	0.219		0.063	0.156	0.063	0.125	0.125	0.094	0.416
P13	0	0.111	0.111		0.111	0.333	0	0.111	0	0.222	
P14	0.167	0.083	0.104	0.104		0.125	0.125	0.104	0.083	0.104	0.769
P14	0.237	0.053	0.079	0.158		0.158	0.053	0.132	0.079	0.053	
P15	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	0.111	0.314
P15	0.308	0	0.231	0.154	0.192		0.038	0.038	0	0.038	
P16	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.111	0.394
P16	0.123	0.053	0.053	0.175	0.211	0.263		0.07	0.035	0.018	
P17	0.111	0.111	0.111	0.111	0.111	0.111	0.111		0.111	0.111	0.564
P17	0.17	0.094	0.057	0.057	0.208	0.094	0.057		0.113	0.151	
P18	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		0.2	0.802
P18	0.065	0.129	0	0.129	0.129	0.032	0.097	0.129		0.29	
P19	0.146	0.146	0.073	0.073	0.073	0.073	0.073	0.146	0.195		0.936
P19	0.135	0.162	0.054	0.054	0.081	0.054	0.041	0.216	0.203		