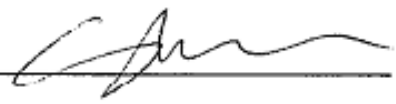
APPROVAL SHEET

Title: Effects Of Group Interaction On Community Based Information Sharing

Name of Candidate: Ajinkya Anant Gaurkar

Master of Science, 2018

Thesis and Abstract Approved: 

Dr. Shimei Pan,

Professor, Department of Information Systems

Date Approved: 11/29/2018

NOTE: *The Approval Sheet with the original signature must accompany the thesis or dissertation. No terminal punctuation is to be used.

ABSTRACT

Title: **Effects Of group interaction On community-based information sharing**

Ajinkya Anant Gaurkar, Master of Science, 2018

Directed By: Dr. Shimei Pan, Professor, Department of Information Systems

The success of community-based information sharing critically depends on the retention and engagement of its members. In this study, we study the impact of natural groups on user engagement and retention in community-based information sharing. To do that, a novel mobile application was developed to capture the activities of users in real time college campus event sharing. We also conducted a user study to exam the impact of group affiliation and group size on user retention, user engagement, and overall user satisfaction. The study results indicate that participants affiliated with a natural group are more likely to engage in using the application longer, using it more frequently and more satisfied with the application as compared to independent individuals who are not members of any natural groups. The study results also suggest that the size of groups could impact user engagement and retention. The practical and theoretical implications of this research are also discussed.

EFFECTS OF GROUP INTERACTION ON COMMUNITY BASED INFORMATION SHARING

By

Ajinkya Anant Gaurkar

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, Baltimore County, in partial fulfillment
of the requirements for the degree of
Master of Science
2018

ACKNOWLEDGEMENTS

I would like to thank my primary thesis advisor, Dr. Shimei Pan, and my committee members, Dr. Zhiyuan Chen, and Dr. Ravi Kuber for their guidance during this enriching experience.

I would also like to thank Siddharth Ashok Patro, Neha Reddy Chandupatla, for managing and maintaining the mobile backend services.

Lastly, I would like to thank all my friends, family, and loved ones who have never wavered in their support.

I could not have completed my work without your strong belief and warm blessings.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Community-based information sharing (CBIS) is any voluntary basis information sharing act appropriating positive social values associated with common understandings of sharing, such as community, generosity, shared values of cooperation, and participation. While there has been a recent surge of interest in sharing, conceptual gaps remain. The success of any applications using the theory of CBIS depends on factors like user engagement and retention. One of the challenges faced is that there are no financial rewards associated with sharing information and the only motivation is that act of exchanging and sharing information has a positive effect on collaborative information exchange. (Reijo Savolainen, 2017). For instance, information sharing on criminal activities in a locality could benefit the community and help law enforcement to curb such activities. There is a wealth of information and research on collaborative information in organizations. But research on what makes individuals use and retain themselves in such voluntary application is elaborated in this study.

This study includes natural groups formed by participants who have real world social ties (e.g., friends and families). In the study, individuals with social ties sign in together as a group to use a CBIS application. By comparing participants signing as individuals versus those signing in as members of a natural group, we aim to study the impact of group affiliation and size on user engagement and retention. The main contribution of this research is to study the impact of natural groups on user retention and engagement and verify whether communications between members of a natural group, some of which are real world interactions not recorded by our CBIS application, can help motivate users to interact with a CBIS application more frequently and remain active longer. The results of our study suggest that engaging members of a natural group together in CBIS applications may provide additional incentive to encourage and improve user retention and engagement.

1.1 Structure of thesis

The next chapter defines the concept “natural group” in the study followed by how user engagement is being studied. The effect of user retention is also elaborated. In addition to this, we understand how the factors like group size a contributing factor in this study is. The thesis is structured starting with background study on user engagement, retention and the effects group size on them. They are used to motivate the main hypotheses and research questions in the study. Then we move to explain the CBIS mobile application design. After that, we discuss the design of the user study to experiment and analyze our hypothesis followed by results, discussion, and future work.

2. RELATED WORK AND RESEARCH QUESTIONS

In this section, we first summarize prior literature regarding community-based information systems, users’ motivation to contribute and the types of incentive users receive, combined with literature study on group collaboration and engagement.

Starting with the motivation, Tanja Aitamurto et al. (2017) examined the motivations and expectations of participants in a community information sharing environment and differentiate how different kind of motivation works, i.e. Intrinsic and Extrinsic. Further, Mokter Hossain (2012) explained the difference between various types of incentives such as, financial, social and organizational incentives and their impact on user engagement. Also “Group collaboration” was also found to improve user engagement in online platforms as mentioned by Yujin Lim et al. (2012). Finally, Anne Ekholm (2012) found the importance of group communication in a mobile environment and explained the design elements that are important to consider while designing group communication platform. It also sheds light on issues that arise during designing a system that supports group communication.

Drawing from the above research work, we have learned that intrinsic and extrinsic motivation coupled with appropriate incentives are important for the success of any community-based information sharing platform. Also, group collaboration could help improve user engagement and retention in online platforms. We then extend the findings from the above-mentioned literature into this research. Our contribution helps bridge the research gap in the area of collaborative information exchange by studying the effect of “Natural Groups” in a Community based information sharing environment. Also understanding how various group factors like group size and affiliation impact user retention, engagement, and satisfaction among users.

2.1 Natural Groups in CBIS

Our experiment focuses on groups formed naturally and consists of people who are familiar with each other. There could exist back channel communications among them, meaning, the participants interact in different real-world social settings in addition to this application. (e.g., friends or family members). Extensive research has been done in the context of collaborative information exchange at workplaces (Danny Yang Chang Ho, 2011) and information sharing in professional communities (Li-Wen Chuang et al, 2017). Then the use of information sharing among students is studied in (Sami Sifi et al, 2014) which employs a collaborative social platform to enhance opportunistic collaboration between students. But the impact of natural groups on user engagement and retention in CBIS applications has not been studied extensively before. This research provides a chance to explore how information sharing platforms can be used in an informal setting by a group of people with natural real-world social ties.

2.2 User Engagement in CBIS

After understanding the environment of the study, we next emphasize the use of information sharing in a voluntary environment to test the impact of group interaction. We now have more opportunities to cooperate with a community or build a virtual team to mutually help one another using information sharing technology. Chuang et al. (2016) found that using self-presentation as motivation, and finally, social network as a factor of collaborative information exchange we can use social media to improve collaborative information exchange. This study takes the theory of group collaboration in social media as the main core to explore how to use information sharing applications conduct to collaborative information exchange. Hence, leading to the hypothesis ***H1:*** *User engagement with the application is improved when users are associated with a group. The* ***null hypothesis:*** *in terms of engagement level, there is no difference between the participants who sign in as individuals and those who sign in as members of a natural group.* **User engagement** in this research refers to how frequently a user interacts with our application and its features.

2.3 User retention in CBIS

Motivation to do good is a positive way to retain users for voluntary applications. Ordinary individuals, however, are reluctant to participate and share information due to a lack of enough incentives. (Xinglin Zhang, 2016) summarized the main types of incentive methods for information sharing as (1) entertainment (2) mutual-benefit (3) monetary incentives. Since our application is voluntary, in additional to giving virtual rewards like weekly appreciation, points given for actions made on the application and giving credit to substantial performance, we also want to study whether encouraging users with natural real-world social ties to sign in as a group can help improve retention. This leads us to hypothesis H2.

H2: *User retention rate is improved when users are engaged in natural groups.* ***The null hypothesis:*** *in terms of retention rate, there is no difference between participants who sign in as individuals and those who sign in as a member of a natural group.*

User retention in our application refers to the ability of our application to retain its users over some period. It begins with the first contact a user has with the application and continues throughout the entire lifetime of the relationship. Here we define Retention rate as follows

R = (E/S) X 100 (Salesforce landmark,2018)

Here R: retention rate in percentage

E: No of users at the end

S: No of users at the start

2.4 Impact of group size in CBIS

The relationship between group size and voluntary information solving is complex. Theoretically, in a problem-solving environment, a small group of researchers has argued that performance should increase monotonically with group size because larger groups possess more knowledge from which to generate good solutions (Steiner 1966). However, the positive effects of group size on performance are often reduced or eliminated by group inefficiencies (Hill 1982). On average, groups only perform as well as their best member and additional group members add relatively little. At worst, groups perform worse than individuals working on their own. Crowds are assumed to benefit from their large size and sidestep group inefficiencies, but studies on the relationship between crowd size and performance are mixed. Larger crowds have been associated with better performance, but only under certain conditions (Kittur and Kraut 2008). Mixed empirical findings in small groups showed the relationship between group size and performance is complicated and context-dependent, requiring more complex theoretical models. However, research on the voluntary environment is different because of the nature of the

information shared. In our study, we aim to understand how group size is affecting retention rate in a voluntary information sharing environment. This led to two new research question

H3a: Group size affects user retention rate. The **null hypothesis**: the retention rates for groups of different sizes are the same.

H3b: Group size affects user engagement as measured by a users' in-app activities. The **null hypothesis**: the use engagement levels for groups of different sizes are the same.

2.5 User satisfaction in CBIS

Information sharing in group initiative creates an experience that may or may not foster engagement and improve user satisfaction. Our research checks the effect of H1-H3 on user experience to understand if the process creates a sense of satisfaction in using the information sharing application. Leading to the following hypothesis

H4a: Being part of a group impacts user satisfaction levels. The **null hypothesis**: the user satisfaction levels for participants who sign up as individuals are the same as those who sign up as members of natural groups.

H4b: Group size affects user satisfaction levels. The **null hypothesis**: users in groups of different sizes have the same satisfaction level.

Based on our hypothesis, we have designed and implemented our proposed CBIS application that would help us collect necessary information and study user behavior. Study design and implementation have been discussed in the next chapter.

2.6 Summary of Research Questions

Hypotheses	**Null Hypothesis(H_0)**
H1: User engagement improves when users are associated with a natural group	User engagement levels are the same regardless whether the user is a part of a natural group or not.

H2 User retention rate improves when users are associated with a natural group	User retention rates are the same regardless of whether the user is a part of a natural group or not.
H3a: Group size affects user retention rate	User retention rates are the same regardless of the sizes of their associated natural groups.
H3b: Group size affects user engagement	User engagement levels are the same regardless of the sizes of their associated natural groups.
H4a Whether a user belongs to a natural group or not affects his/her overall satisfaction levels	User satisfaction levels are the same regardless of whether he/she belongs to a natural group or not.
H4b: The size of a group impacts its members' satisfaction levels	User satisfaction levels are the same regardless of the sizes of their associated natural groups.

3. CBIS APPLICATION DESIGN

Users expect a lot from mobile apps today, and the expectations are just getting higher. Meeting those expectations is a fairly difficult task to make the system useful, relevant, and valuable for users. Designing a similar system for research purpose is equally complicated. In this research, we have proposed and developed a design concept for a mobile application called "*Snapeve*", that would involve many participants to contribute to the application. The data collected by this application would then be studied and analyzed to understand the validity of our hypothesis and the feasibility of this design for a large-scale implementation. There are many things to consider when designing an application for mobile, especially in a community-based information sharing setting, where a little amount of knowledge gain requires a significantly large amount user contribution. Various factors are important and worth considering while designing such type of applications. Snapeve is developed with the following design criteria in mind in order to encourage maximum user interaction (Nick Babich, 2018).

3.1 Design criteria

A. Minimizing Cognitive Load

Cognitive Load refers to "*The total amount of mental effort that is required to complete a task involving processing of information"*. Snapeve involves various techniques and features that help reduce cognitive load on users, such as breaking the task into many pieces, decluttering the information and visual instructions while performing new activities. In addition, forming user groups to divide a tasks' load would be an effective way to reduce cognitive load on users.

B. Designing a good onboarding experience

In the context of the mobile UX, delivering an excellent onboarding experience is the foundation for retaining users. The goal of onboarding is to show the value your app provides. Among the many strategies for onboarding, contextual onboarding is especially effective. Contextual onboarding means that instructions are provided only when the user needs them. Snapeve uses interactive onboarding screens to help users understand what they will be doing once they are onboard and give them a visual clue of what the app will look like.

Some user might be confident enough about their understanding of technology and user experience and would feel the snapshots presented during onboarding are time consuming. To overcome this situation, a simple option to "skip" the snapshot's screens and direct the user to a meaningful page would help maintain user's interest and make the user less frustrated.

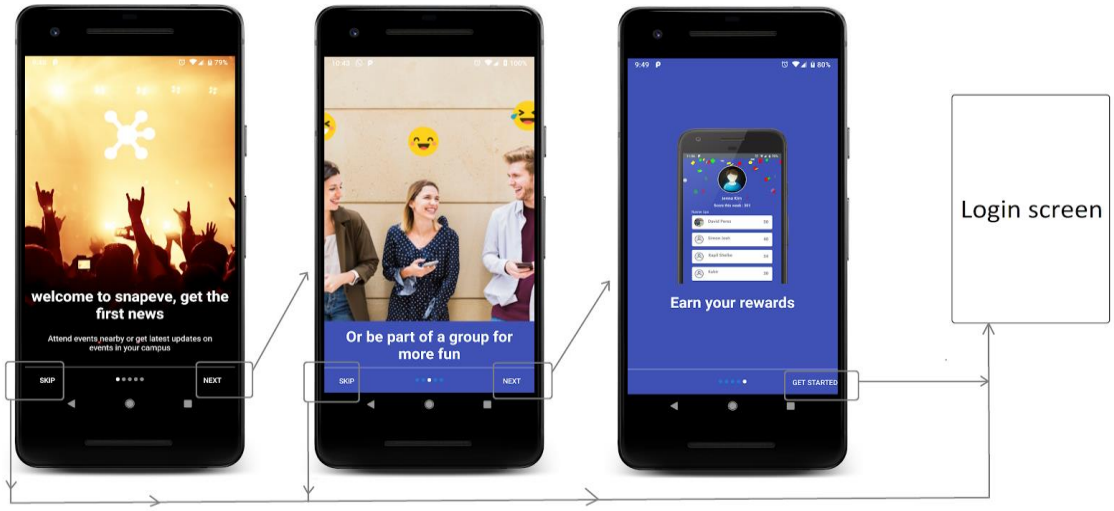


Fig 3.1 Onboarding process with skip options

C. Notifying users about activities happening in the group

Notification plays an important role in improving user engagement. Users tend to lose track of what applications they have installed on the phone. Notification can remind them that they have a privilege of performing some sort of actions into an app. CBIS is one such concept where user interaction is needed the most, notification is an effective channel to connect the user to the system when needed. As the notification matrix shows, CBIS events should have real-time notification display mechanism due to high urgency. (Rasha el Stohy et al, 2016) This paper focuses on Push notification the most to improve real-time communication and user engagement.

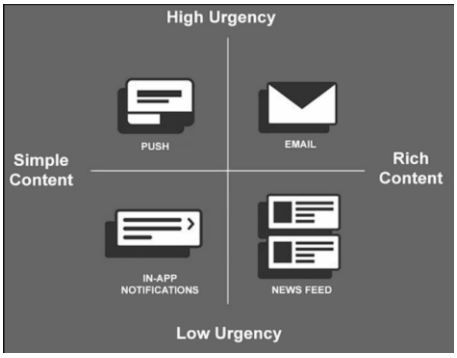


Fig 3.2 shows the matrix of different message pushing technologies based on their usage

D. Providing feedback to every contribution made by user

Users might lose interest in doing similar tasks again and again. This is one big issue with collaborative information exchange platforms where a simple task is repeated multiple times by numerous users or by the same user. Providing dynamic feedback wherever possible ensures the longevity of the application and would improve retention of the users. Snapeve uses interactive feedback messages to keep users engaged.

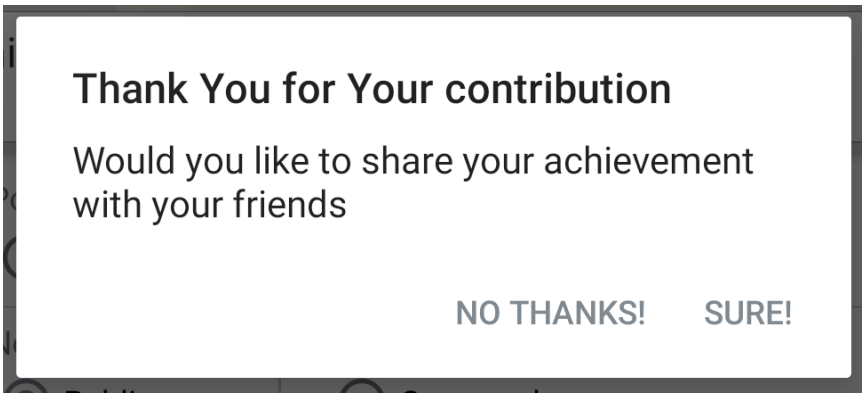


Fig 3.3 Shows appreciation message after an event is posted

E. Prioritizing navigation options

Prioritizing navigation based on the way users interact with your app and assigning different priority levels (high, medium, low) to common user tasks is important while designing applications. Giving prominence in the UI to paths and destinations with high priority levels and frequent use can help users logically map the structure of the app into their brain. And those paths should also be used to define navigation in the app. Organizing information structure in a way that requires a minimum number of taps, swipes, and screens will help the user to get used to the app more quickly.

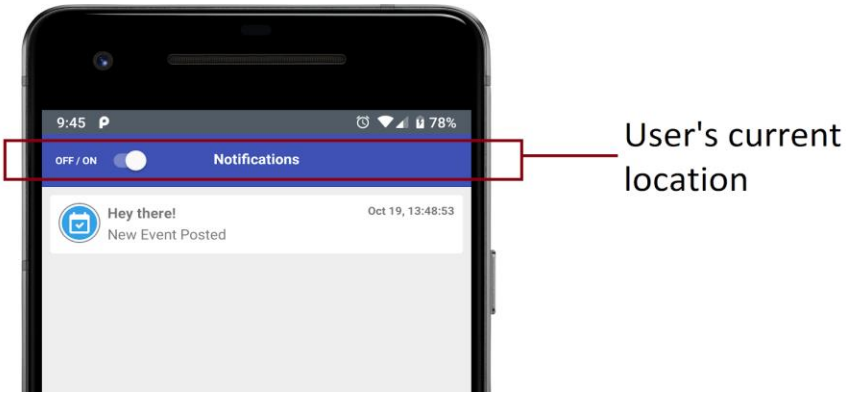


Fig 3.4 The screen header shows current location of the user in the application

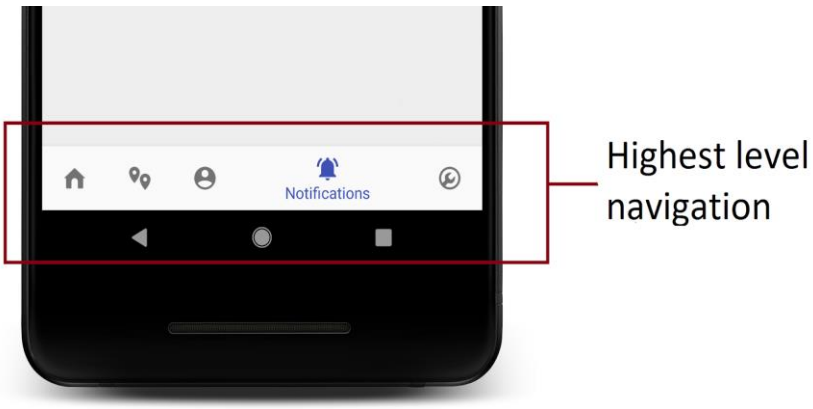


Fig 3.5 The five important navigation actions that are visible globally in the app. These actions are expected to be used most frequently.

F. Displaying meaningful error messages

Errors occur when people use applications. Sometimes, they happen due to the user's mistake and sometimes because the app fails. Whatever is the reason, these exceptional situations and how they are handled have a big impact on user engagement. Bad error handling along with less relevant messages can raise users' frustration and could be the reason why users opt out of any mobile platform. Assuming users as tech-savvy is not a good approach. It is always better to inform users of what's wrong in plain language. Each error message should indicate:

- what went wrong and possibly why,
- what's the next step the user should take to handle the situation.

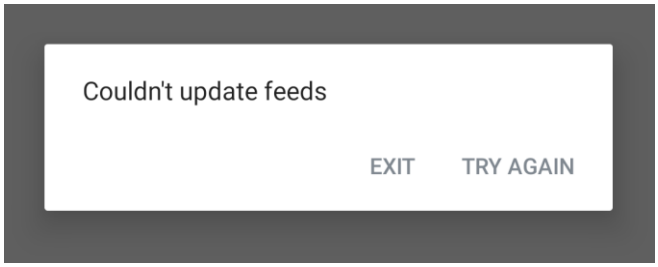


Fig 3.6 Shows the reason why data cannot be displayed on the screen with two possible actions

3.2 Essential features for user activity

To understand the effect of group on user engagement and user retention, we have developed some features that required participants to perform certain actions either as an individual or as a part of a group. Group features are based on the features for individual users but are an extension to them with more options to engage with group members. Features for individual users are listed below followed by group features in section 3.2.3

A. Posting an event

Snapeve (“Snap an event”) is based on an event sharing idea where participants can use their smartphone camera to take pictures and share it through the Snapeve mobile application. This is the most important feature and is the starting point of almost every use case cycle in the app

B. Verifying or spamming the event

Once a post has been made by a user, all other users or the group members of the poster get notified about the post (notification depends upon the “Notify all” choice made by the poster). Other users have an option to either verify a post as a legit by hitting “Verify” icon or mark it as “Spam” if they find it as inappropriate, false or misleading.

C. Commenting

Apart from verifying or spamming, users also have an option to “Comment” on the post. Comments can be made by any user to whom the post is visible. Comments can be added to a post which is considered generic and undirected to any user, but a comment can also be directed to a specific user, which defaults to group members. This is made with an intention to improve communication.

D. Map view of live events

Every post made in the application also has a location parameter to provide a user with the information about where the event is happening. While making a post, a user can either select a geolocation on the map or select a location from a list of commonly found locations from a picklist or choose the current location (Default option) if the user is posting live event. All these options collect location coordinates from the user and display it in a dedicated map’s section in the application. This feature gives users the ability to quickly check the locations of the events without struggling to find the location in a vertical list of events on the dashboard.

E. Sending invites to users to attend an event

Every event that is displayed on a user’s screen has an option to mark users’ attendance status. Various statuses such as “Attending”, “Interested”, “Not interested” are available for users to mark. In addition to a user showing interest in an event and marking the status according to the interest, other users can also request a user to attend an event. This feature is called as “Sending invites” of events. If a user sends out an invite to one or multiple users, other users get a notification with the event details and a message requesting to attend (or checking) an event. Sending invites would increase collaboration amongst users and popularize an event that is expected to engage more and more users into the application.

F. Rewards for user contribution

Users need appreciation for their efforts. Appreciation in the form of monetary benefit or virtual money is an easy way to engage users into the system. But in our case providing monetary benefits is not the option or choice, rather intrinsic motivation to post events that would benefit the community is the driving force behind this application. Hence, users are rewarded with virtual points (with no monetary value) to keep them engaged in the system. Tasks like posting an event, verifying/spamming, commenting gives users' virtual points that would show up in their account in the application. In addition to the rewards every user earns, there are rewards for users who are part of a group. These rewards are called Group rewards and are calculated differently which have no effect on individual earnings and vice versa. To earn group rewards, users must perform group activities and the rewards are shared between group members to improve collaboration and a sense of group reputation.

G. Leaderboard to show statistics of other users

Users need a reference to gauge their earnings, hence "Leaderboard" gives them a reference of how other users are performing in the application and how much more or fewer efforts a user has to invest to compete with others. A leaderboard is an effective method to improve participation in any competition, the same concept applies in mobile application as well where the prime motto of the application is to gain users attention and their inputs. Snapeve has a leaderboard for both individual users and group users (or group leaderboard). Group leaderboard compares rewards and earning of one group to other.

H. Weekly recognition of users based on reward points earned during each week

Apart from the rewards given to users after every task they perform, there are weekly rewards available that come as a notification to the user's device every weekend. In short, weekly

recognition is a way to appreciate the contribution of users on a weekly basis. These rewards are based on the rewards users earn while using the application but are calculated on weekly basis and the values are reset after every week. This gives chance for other users to perform better and earn more rewards and gain recognition for a week if their total rewards are significantly low compared to total rewards of the top scorer within in the application.

I. User profile page to display status and reward points

The above-mentioned reward system is accessible from user profile page which has options to switch between the following:

- User rewards and user contribution
- Group rewards (if the user is associated with the group) and group member list and their contribution
- Leaderboard list of individual and group users
- Weekly rewards (if weekly rewards are declared for that week)

J. Notifications for updates to the events

Push notification is an effective channel through which users can be notified about the latest updates of any community based mobile application. Snapeve uses notifications to engage users into newly posted events, react to others' comments and stay updated with the group's activity using group notifications. Notifications can be sometimes annoying, hence Snapeve provides options to users to manage notification according to their convenience. They can choose to stop or get notified about group activities, manage notifications on comments or invites and schedule them at preference.

3.3 Group features to support group member interaction

This paper focuses on how in-group communication and group size affect the following features.

A. Post an event on behalf of the group:

Being part of a group inculcates users to post sensible content because they are mindful of their group's reputation and would want to bring value to their group. At the same time, it improves the group's incentives if users make a post on behalf of the group.

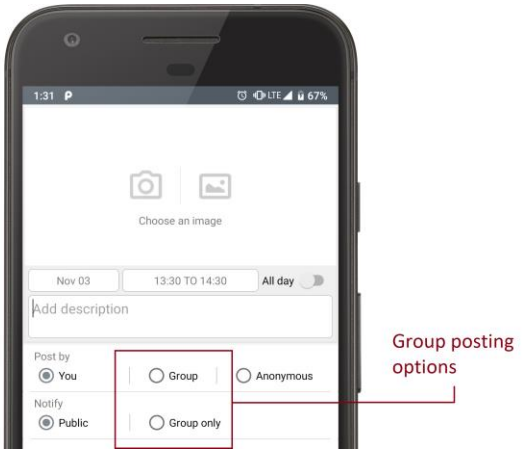


Fig 3.7: Figure shows options to post event on behalf of group

B. Sending event invite to group member (invite defaults to group member first)

Snapeve allows a user to send invites for an event that is posted in the app. Users can search from a list of other users to send invites for an event, but the list emphasizes on group members and shows their usernames exclusively adjacent to the list that shows usernames for all the users in the system. This allows group members to promote events to other members of the group.

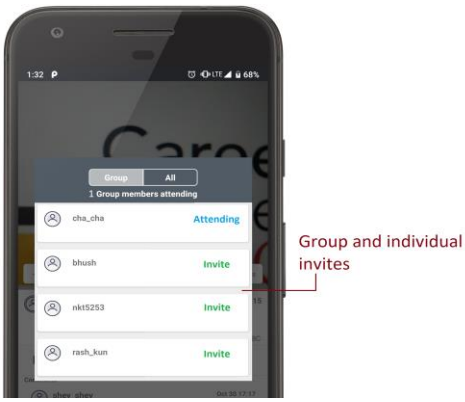


Fig 3.8 Figure shows event invites that defaults to group people

C. Multilevel group management privileges

The group management provides group administrator with the highest privileges and the bonding of a group would garner the privileges grants to other members of the group. The app includes various levels privileges highest being able to make other group members as an admin, the ability to add or remove new members to the group with the most basic privilege being able to see and make posts. The above group features incorporate the essence of group's bonding and its positive effects on the effectiveness of a group. In simple terms, higher the privilege stronger the bonding between group members

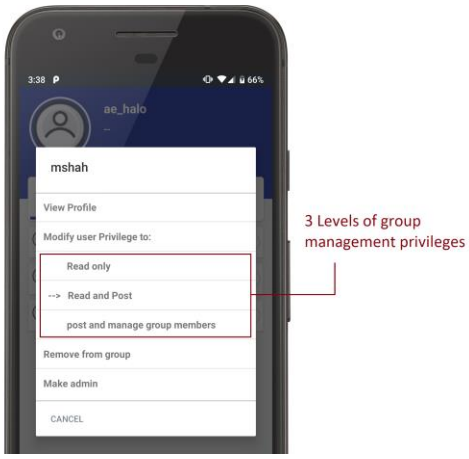


Fig 3.9 Figure shows options for admin to grant privileges to group members

D. Group rewards

Each user activity receives some rewards as virtual points. Activities done on behalf of group gains more points for the group. One way to encourage group interaction is to give users more benefits if they participate in the app as a member of a group and do activities on the group's behalf. Group activities, when compared to individual activities, shares similar reward system but rewards in group settings are usually higher and are more prominent on group dashboard, which gives the users a sense of group collaboration and group reputation.

Weekly group scores and rewards

Fig 3.10 Figure shows weekly rewards and team members' contribution for a given time frame

E. Directed Group Comments to improve group communication

Comments can be made in the application in two ways. First is the "Directed comments" which is made on a target user. This enables users to communicate in the app in a one-to-one manner. Second, undirected comments, where comments are generalized for a bigger target audience and do not have a target user in its content.

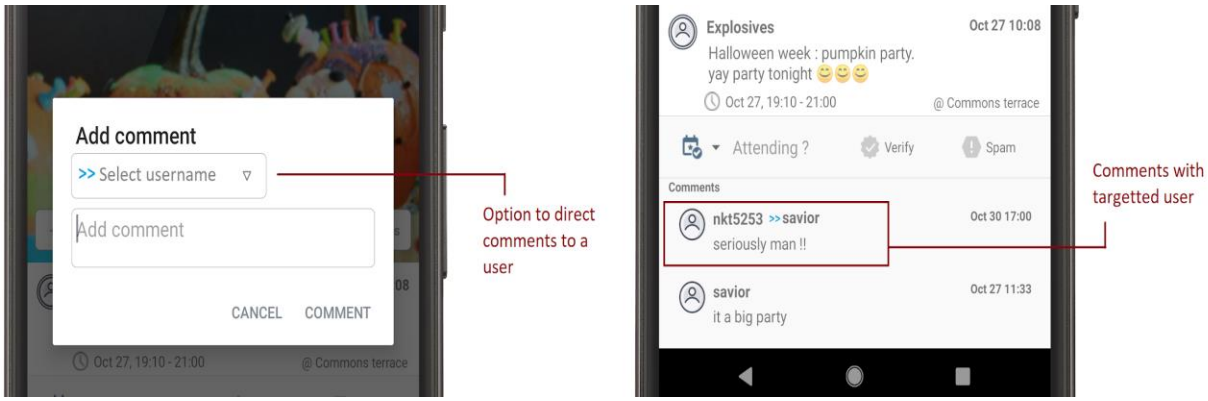


Fig 3.11 Shows option to direct comments to a user

F. Dedicated notification for group related activities

The notifications released for every single activity performed by a certain member of a group can be controlled by the actor. In doing so, a participant who is part of a group may notify every member of their group of his/her activities and thereby increase group engagement. Also, they have the option to notify every one of their actions and a receiver may choose to get notifications of only their group or everyone. This feature helps user engage in the activities of their fellow users in the app.

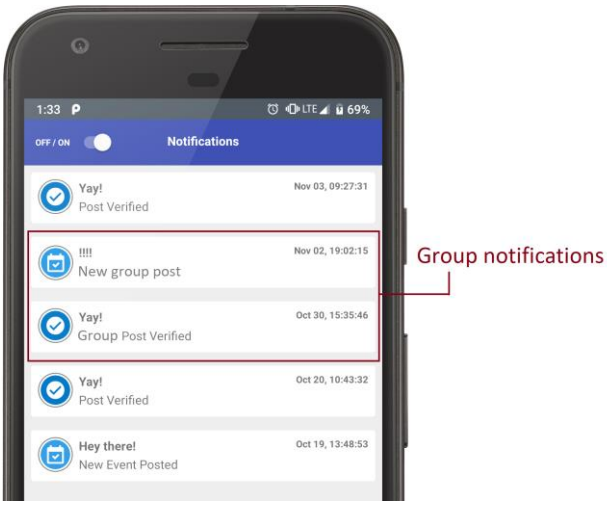


Fig 3.12: Figure shows dedicated notifications for group activities

F. Group dashboard

Snapeve displays comparative statistics of an individual's performance in relation to said individuals' performance as a part of a group. This is a form of motivation to the users by providing a comprehensive view of their performance in the application till date. This enables users in the application to develop a sense of competition in order to improve user engagement.

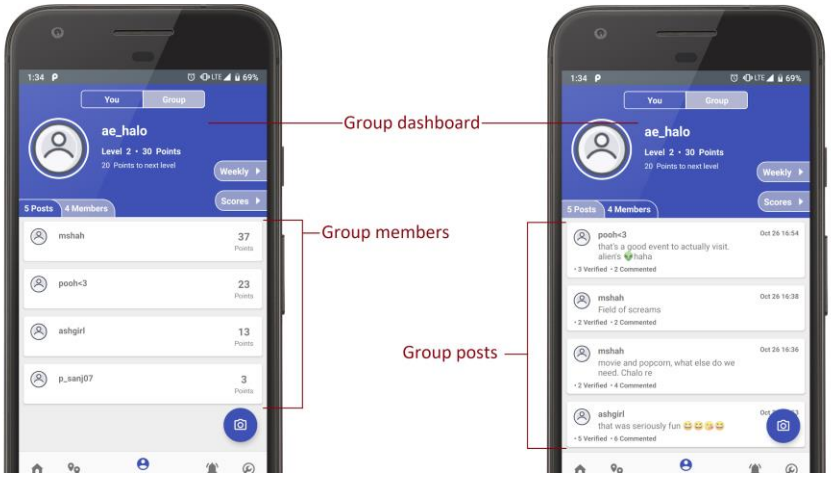


Fig 3.13: Figure shows group dashboard with group scores, members and total group contribution

3.4 Technical implementation

Snapeve is an Android-based mobile application. Given the scope of this research and budget constraints, Android is chosen as the platform for application development and use. The architecture below lists the vital components of the system and it shows how are they connected to each other. Details about each component and their implementation is explained in the following subsections.

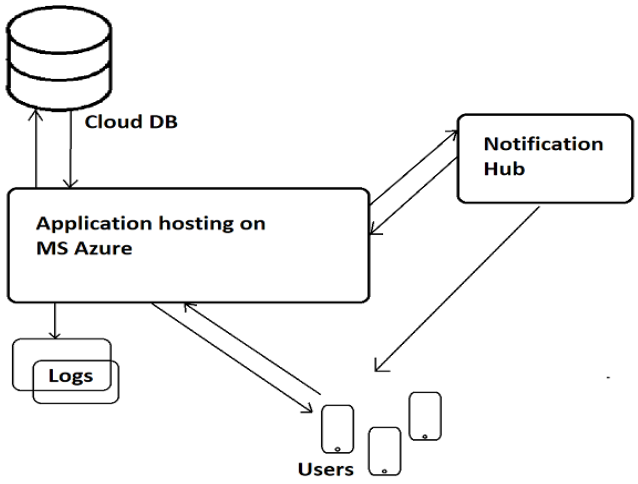


Fig 3.14: Shows the Snapeve architecture and its vital components

3.4.1 Application development

A. Concept & wireframing

Based on the research requirements, the concept for the application is developed. Features were designed to improve group participation. Wireframing is done to convert the design ideas into tangible screen components. Wireframing ensured that all the designed features matched the requirement, which can then be transformed into real application components.

B. Local Environment setup

The application is designed to work on Android operating systems, which required the setup of Android studio and other frameworks for mobile app development. Studying the market share of the operating system's versions is an important step before code implementation. Based on the required versions and their share, all necessary frameworks were integrated into the local development environment. The Android version 4.4 to 8.1 includes over 90% of the android users in the world and users running android below v4.4 can be excluded from the user base of

the app as their device is most likely to become obsolete and would stop supporting any new application released in the market. And on contrary, the users who have the version of the latest and greatest operating system can also be excluded from the current version of the app as their number is significantly lower to consider including as a supported version for Snapeve. Support to such users can be given in the later releases of this application. Moreover, our target demographic is students who tend to use mobile operating systems ranging between the above-mentioned range.

C. Cloud data storage setup

The whole setup is dependent on Microsoft Azure Cloud data storage to maintain and store application data. Due to the wide scope and unanticipated user base of the application, cloud infrastructure is the most feasible option, as it provides a scalable development platform for application development and deployment. In addition, all other required frameworks such as application analytics, session management, and media storage come pre-equipped with cloud infrastructure.

D. Libraries

Various third-party open source libraries were used during the application development. Many features in the application were graphic intensive and required writing complex code. Such type of codes could be made available by using open source graphic libraries. In addition, non-native frameworks such as in-app survey, animations, custom text formats, image loader were implemented using open source libraries found on GitHub which were mostly sourced by MIT, (Qualaroo marketers guide,2018).

E. Notification system

This application is heavily dependent on the push notification system to engage users into the app. Users received real-time notification about the activities going on in their vicinity. Firebase push notification system is used for this purpose. Code package provided by Firebase is integrated into the application to receive notifications and to enable a communication channel between mobile device and notification hub.

F. Application Testing

Testing is a major and important part of application development. Mobile development possesses different challenges when it comes to testing the application. Mobile operating systems have various versions as opposed to a computer-based operating system. Mobile devices also vary in screen sizes and screen display density. Hence the scope of testing increases exponentially when it comes to test mobile applications. To solve this problem, “Mobile device farm” is used. It is a cloud-based service that involves virtual mobile devices of all varieties ranging from various screen sizes and all operating systems and their versions that are currently active in the world.

G. Data logging system

In order to study user behavior and usage pattern, user activity must be logged, and every action of the user is noted down in the system. In this application, logs are collected in two ways:

I. First, data logs that are calculated from users’ in-app activities that are visible to other users. Examples include posts made in the current week, comments made or the number of comments after a post is made by another user. These type of data logs can be obtained

from the data which gets collected on the application server and can be queried and extracted on the ad-hoc basis.

II. Second, the frequency of visits of a user into the app, time spent on a screen or sequence of screens a user goes through during his or her active session. These data logs are not available real time. Instead, they are locally stored on a user's device whenever the user uses the app and the data is pushed to the server in batches after a specific time interval. This kind of batch processing is implemented to minimize the computation load on the application as the user uses the app, and once the app goes into the background the batch processing starts according to the predetermined schedule.

3.4.2 Snapeve Deployment

Based on the study about system design, we implemented the features in the application that facilitated community-based information sharing and improve group collaboration. The application is available on G*oogle Playstore* for free to download for participants of the study

4. USER STUDY DESIGN

The following content explains the study design that will allow us to understand user behavior and draw effective conclusions to the hypotheses.

But Before we collect any data, we must complete the training for research study conducted on human subject. Paperwork for a study is filed which is decided to be under this criteria- Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies. (NOTE: Some research in this category may be exempt from the HHS regulations for the

protection of human subjects. 45 CFR 46.101(b)(2) and (b)(3). This listing refers only to research that is not exempt.)

In context of this research, the study will be conducted without collecting any personal identifiers from the participants and they will only use the Snapeve application upon providing their consent. This would be informing them of their rights and the data being collected has no reference to their personal identity. IRB required us to provide information on the data storage, access permissions, and duration of the study along with the kind of investigation that would be performed on the data after the study. A copy of the questionnaire is sent in advance to screen the type of information participants would be required to give during the feedback process.

We move to the user study to implement the above architecture in order to study our hypotheses. To do that, we recruit participants who can contribute to our study. It follows the following process:

4.1 Recruitment

The recruitment process for this research study began with recruitment flier distribution. (Kubicek, K., Robles, M. 2016) Explains how fliers are one of the most effective media to recruit participants. Since this study is based on a college campus population, distributing fliers is a feasible option. In order to collect more participants, the campus population is also emailed regarding the research study using the public email distribution list. Digital fliers were attached in the emails with an intention to provide fliers to those who didn't get a chance to see the physical fliers. At last, various event boards on campus were approached to spread the details about this study.

Once the potential participants decide to participate in the study, they are expected to enroll via the participation link mentioned in the flier. They are presented with the choice of participating

as an individual or as a group. Also, they are asked to provide 3 distinct aliases for their profile in case there are conflicts between the alias chosen by multiple users. If a participant decides to participate as a part of a group, he or she is asked to provide an alias of the group as well as the aliases of the group members. The choice of group members is left with the participants since our target is to form naturally occurring groups. To facilitate the study of the impact of group size, we limited the group size to a maximum of 5 and minimum of 2 to prevent the formation of a small number of large groups.

Following the participant onboarding process completion, the principal investigator creates accounts with user provided aliases for every user and default passwords, which are expected to be changed by all participants during their first login attempt. The Snapeve application is coded in such a way that a user needs to go through a password reset process if it is a first-time login attempt. Since this research study is based on human subjects under section 45 CFR 46.101(b)(2) and (b)(3), according to the IRB guidelines, a person who is not related to the study and have not read/accepted the user consent form, is not supposed to sign up or login into the application. Hence the account creation is isolated and only done by the principal investigator. Participants are provided a link to download the app which is hosted on *Google Playstore* for free.

Moreover, like other social media applications, before the formal study, Snapeve has populated with campus event data from the pilot study, student boards, and other event organizing boards. This is done to present some information to the participants upon initial login and avoid blank spots in the application. Also, during the onboarding process, participants were given instruction to answer a questionnaire at the end of the study. Notification to Questionnaire is pushed to their smartphone and it is embedded in-app for easy access. This is done in order to reduce extra efforts taken by participants to answer the questionnaire which in turn encourages more response rate.

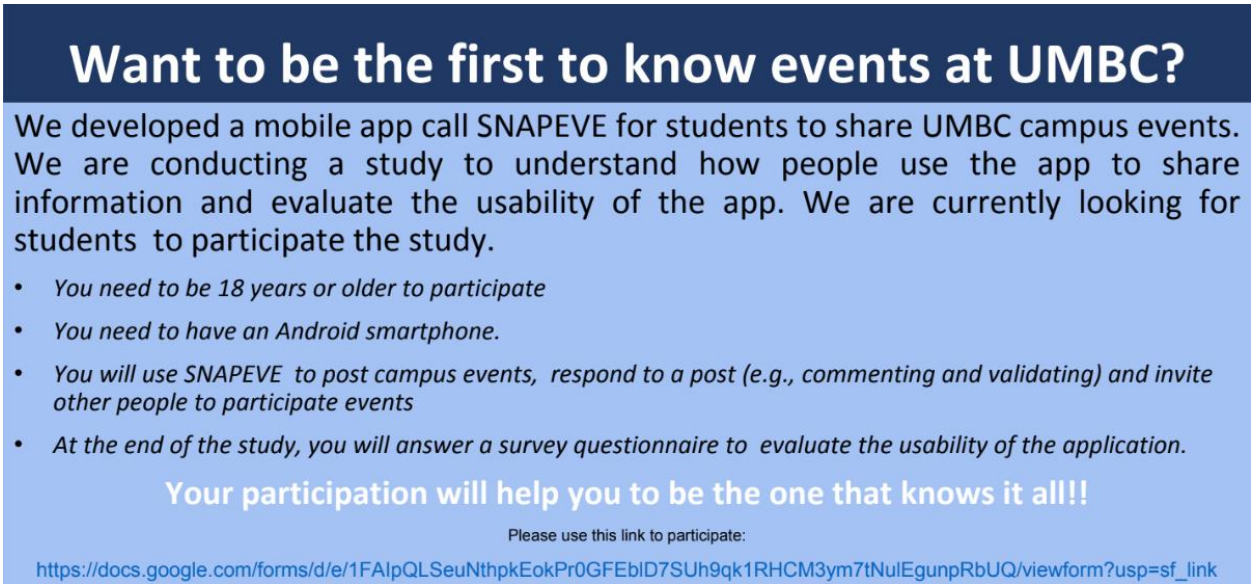


Fig 4.1 Recruitment process - flier for recruitment announcement

4.2 Participants

For this research study, the target user base is the campus population. All the participants were aged 18 or above. After the recruitment, there were 46 participants aboard in total. Among them, there were 21 participants with individual accounts and 25 of them belong to groups. Distributions of participant and group sizes are explained in the table below.

Total Participants	**Group participants**	**Individual participants**	**Total groups**
46	25	21	8

Table 4.1: Participant distribution

Size of group	2	3	4	5
No of Groups	3	2	2	1
Total Participants	6	6	8	5

Table 4.2 Distribution of participants in groups with different sizes

4.3 Evaluation Measures

Once the recruitment process is complete and the participants are aboard, we need to collect data to analyze their behavior. This can be done in various ways. In this research, we are focusing on two specific measures. The first type is the objective measures based on user activity logs automatically collected by the system. Second, are the subjective measures based on the survey questionnaire to help us understand user behavior from a user's standpoint (discussed in section 3.2.4).

4.3.1 Objective Measures

We have designed several measures based on our hypothesis, to measure user engagement, retention in both individual and group settings. They are as follows:

Measure 1: User retention rate

User retention rate represents how many users are retained by the concept or by the design of the application. It is the ratio or number of users remained active in the system to the total number of users with which the application has started in a specific time frame, in our case the time frame is the start and end of the user study.

User retention is calculated using formula **R = (E/S) X 100.** (Salesforce landmark,2018)

Here R: retention in percentage

E: No of users at the end

S: No of users during the start

To study the effect of group on user retention rate we have observed the activity of users where they log in and log out of the application during the research period. If a user logs out the retention rate is considered to decrease but if the same user logs back in after a certain amount of time, the decreased retention rate is neutralized by the login process. Hence the number of login and logout attempts made by a user in the intermediate time is not taken into

consideration, rather only the start and end points are used. Every login activity is denoted with a ‘1’ in the database, similarly ‘-1’ for logout activity, which makes the calculation simpler. Attributes that are considered for user retention analysis are as follows:

Date, User_ID, Grp_No_grp, Login_Value, Lougout_Value

Where “Date”’represents when the activity occurred, “User_id” shows the unique ID of the user, “grp_no_grp” shows the association of a user with any group (if exist), “Login_value” & “Logout_value” represents what action has been taken by the user at that specific date.

By manipulating these 5 attributes we could then study the user retention pattern for individual members and members of the group. Moreover, aggregating values and plotting graphs based on these attributes would render a visual representation of user retention. In addition, we could also perform two sample t-test (Anesthesiol, K. J, 2015) and ANOVA test (Restor Dent Endod, 2014) based on the number of active users per day in both settings to find the significance of patterns in the data.

To study the effect of group size on user retention rate, we have utilized the same concept as of individual vs group user retention, but the comparison here is based on the users belonging to different group sizes. This analysis is based on the users who specifically belong to a group. In our study, we have group sizes ranging from 2 to 5.

Attributes that are considered for user retention analysis on group sizes remain the same as in the previous setting, but the “grp_no_grp” attribute is replaced by the size of the group to which the user belongs

Date, User_id, grp_size, Login_value, Logout_value

Here “Date”’represents when the activity occurred, “User_id” shows the unique ID of the user, “grp_size” shows the size of the group associated with the user, “Login_value” & “Logout_value” represents what action has been taken by the user at that specific date.

Measure 2: User Engagement

In this research, there are various features that are built to analyze user engagement especially for users that are a part of a group. There are two important metrics for measuring engagement, frequency of users checking/visiting a feature and time users spend on a feature. In short "Frequency" and "Duration" are the two key performance indicators. Also, to see the significance in data, we have performed two sample t-test and ANOVA test on the generated results. All these above measures help us derive the results for our third and fourth hypotheses, both related to user engagement. They are as follows:

4.3.1.1 Number of Posts

To measure user engagement from the type of posts made using the Snapeve application following attributes are considered

Date, post_type, user_type

Where "Date" represents date and time when the post was made, "*post_type*" shows what type of posts it is and "*user_type*" tells us about the type of association of a user with the group, if any. Based on these 3 attributes we could study the distribution of posts in the application in various user settings using the following information about posts:

- Posts made by individual users compared to group users
- Comparison of posts made by user belonging to different group sizes

4.3.1.2 Number of Notification Conversions

Attributes that are considered for measuring user engagement from notifications are:

Date, notification, con_notification, grp_notification, con_grp_notification

Where "Date" represents date and time when the notification was pushed to the device, "*notifications*" shows total daily notifications sent to all individual users, "*grp_notification*" shows total daily notifications sent to all group users. "*Con_notification*" and "*con_grp_notification*"

represents how many notifications out of the total notifications were converted to clicks, which can be used as measures of user engagement.

4.3.1.3 Number of Group Feature Visits

To measure user engagement based on user activity in group features such as Group profile, Leaderboard and weekly rewards. we analyzed the usage pattern using in the form of frequency and duration. Following attributes were considered for analysis:

Activiy_code, start_time, end_time

Here, "*Activiy_code"* shows what activity is being performed, *"start_time"* and *"end_time* are the start and end time associated with that activity. Duration of that activity is calculated based on start and end time. Based on the above attributes we are:

- *Comparing users belonging to different group settings based on their visits to group features and duration of their activity. (*Toporek, A. 2017)

4.3.1.4 Number of Comments

Attributes in consideration for measuring user engagement from comments are as follows:

Date, comment, src_user_type, trgt_usr_type

Where "Date" represents date and time when the comment was made, "*comment*" is the content in the comment, "*grp_notification*", "*src_user_type*" and "*trgt_usr_type*" represents if the comment was an individual comment or a directed group comment. Based on these 4 attributes we could study the comments and their contribution in engaging users in the application in various user settings using the following information about comment:

- Comments made by users in individual setting compared to group setting
- Comparing comments made by users belonging to different group sizes

4.3.1.5 Number of Invites

Finally, to measure engagement from the invites sent for the events listed in the posts, following attributes were considered

Date, invite_type, user_type

Where “Date” represents date and time when the post was made, “*invite_type*” shows what type of invite it is and “*user_type*” tells us about the type of association of a user with the group, if any. Based on these 3 attributes we could study the pattern of invites in the application in various user settings using the following information about posts:

- Invites sent to individual user vs group invites

4.3.2 Subjective evaluation

In addition to data analysis, we are analyzing the user responses from the survey questionnaire. These are the self-reporting measures that would provide us with insight of what users are thinking about the system. Survey questionnaire helps us communicate directly with users and understand what their needs and opinions are. We have designed two survey questionnaires for this study.

Individual Participant Questionnaire: General usability questionnaire. It consisted questions that are generic to mobile applications. It is implemented to test the usability of Snapeve. Respondents to this questionnaire would be the participants who participated as individuals.

Group Participant Questionnaire: In addition to all the questions in the individual participant questionnaire, it includes additional questions for measuring the usability of group features that only accessible to group participants.

Open Ended Questions: Open-ended questions are those where users do not provide any standard answers to a question. Open ended question asked in the survey was “*Do you have any comments about Snapeve?”*

Participants in this study were informed before the study commenced, about the survey that they must take at the end of the study. At the end of the study, Snapeve utilized In-app notification systems to remind participants about the survey they need to take by the end of the study. Since the survey is integrated into the application, it is well embedded in the application which they are already using during the period of study. This helps to reduce a user's burden to answer the survey. Additional reminder notifications were sent to the participants every 2 days if a participant did not take the survey. This is expected to improve the response rate.

4.4 Pilot study

The pilot study consisted of 4 users, 2 belonged to a group and other 2 contributed in the individual setting. The aim is to make sure that system implementation, session logs, and main design features work according to its design. We also tested the study flow to make sure it is streamlined. The duration of the pilot study is 1 week

5. RESULTS & DISCUSSION

We performed statistical analysis to derive answers to our main research questions.

5.1 Retention Analysis

The average retention rate of participants was around 69.3%, where, individual participants retention rate was 66.6% compared to group participants with retention rate of 72.0%. Further analysis on group size revealed the retention rates of each group size. We found the retention rate of group size 2,3,4,5 is 83.0%, 83.0%, 75.0%, 40.0% respectively.

Similarly, average retention durations are

Individual users: 15 days

Group users: 15.28 days

Group size 2: 16.66 days

Group size 3: 15.33 days

Group size 4: 15 days

Group size 5: 14 days

The above numbers suggest that group participants are more likely to stay with the application for a longer period. But participants associated with smaller groups tend to remain in the app longer that those in larger groups.

In addition, we also analyze the trend of user retention and visualize them in the following plots, following (Leahy, J. 2004).

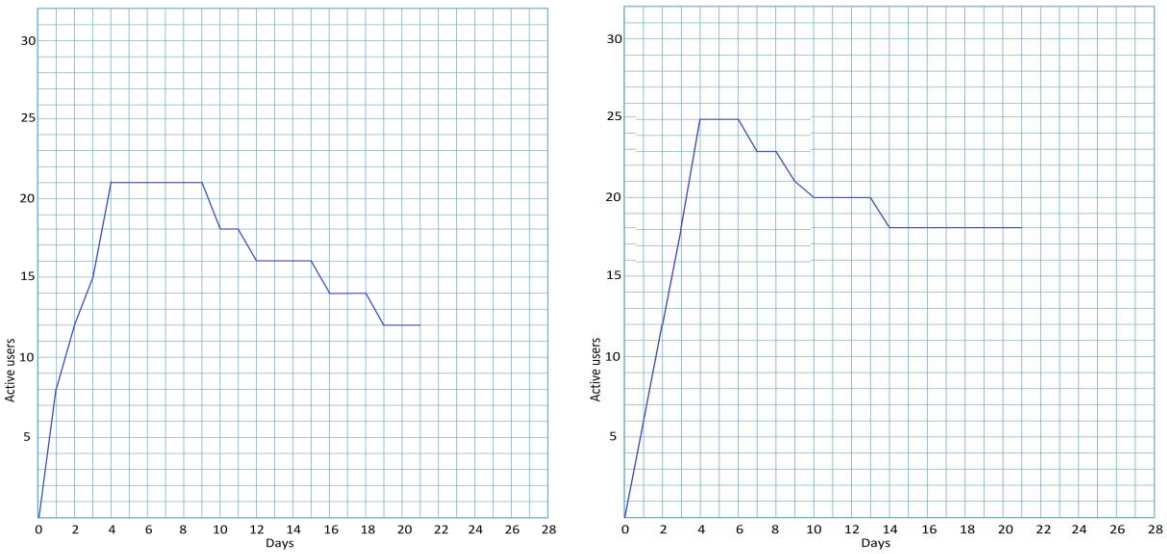


Fig 5.1 shows daily active individual users vs daily active group users

On the above data we performed a two-sample t-test and found out the following results

Attribute	Individual	Groups	T-value	P-value
Mean	15.0	15.28	**2.36969**	**.021358**
Std. dev	3.08	3.10		

Table 5.1 Shows retention for Individual vs group, the difference is found to be significant at $p < .05$.

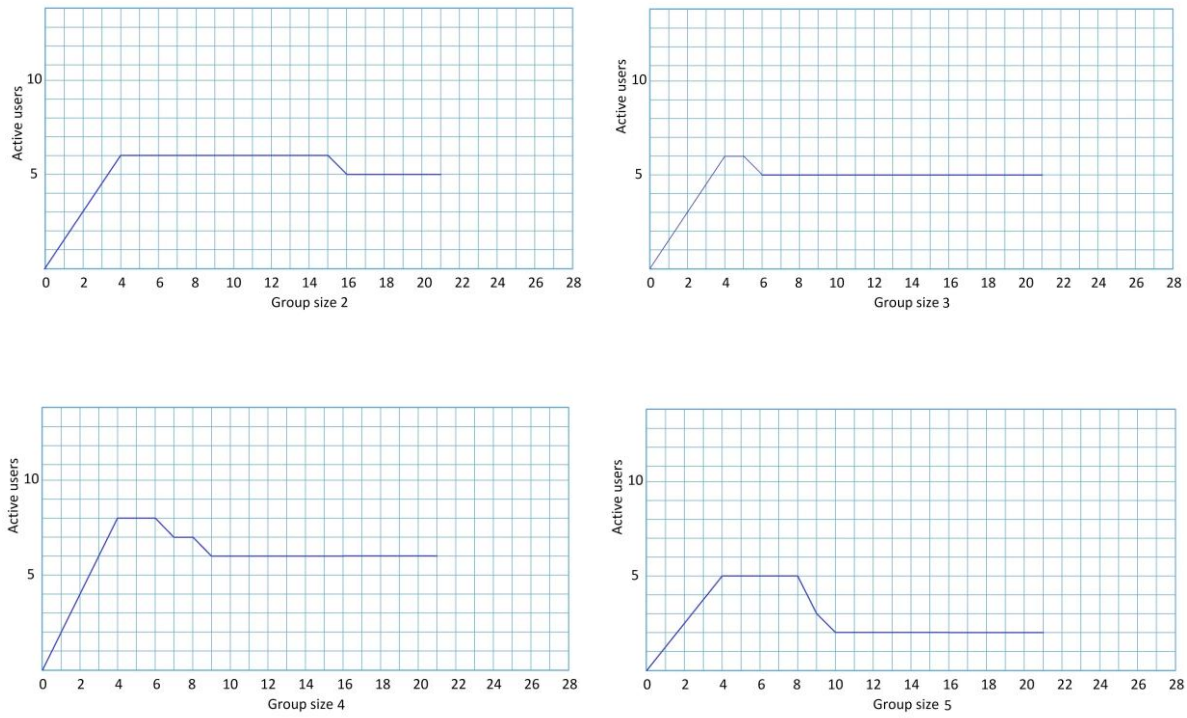


Fig 5.2 shows comparison of active users per day between groups of various sizes

ANOVA test on the above data of 4 different size gives us the following results:

Attribute	Size 2	Size 3	Size 4	Size 5	F-value	P-value
Mean	16.6667	15.3333	15	14	**6.41758**	**.012726**
Std. dev	3.08	3.12	3.2	3.3		

Table 5.2 Shows retention Among various group sizes, results are found to be significant at $p < .05$.

Analyzing the above plots, we can see that individual users stayed longer in the application than group members during the beginning of the study. But, after the initial period, the participation of individual users drops. On the other hand, activities of group users stay high longer. We hypothesize that group members tend to take more time to adjust in the group environment initially. But once a group norm is established, interaction among group members can help sustain retention throughout the study.

In addition to the above analysis, comparing retention between user belonging to groups with different size, we can observe that there is more retention for smaller group sizes compared to larger ones and the difference is statistically significant. There is almost 107% difference between retention rate of smallest and the largest group. Retention rates of other group sizes lie in between the above ranges. Hence the results from this analysis support the hypotheses that both group affiliation and size affect user retention.

5.2 Engagement

User engagement in this research is measured based on users' interactions with main application features. They are discussed in detail below:

5.2.1 Number of Posts

Analyzing the number of posts made by individual users and group users we get the following results:

Total number of posts made in the application during the study period was 49, where individuals have 23 posts and group users have 26.

Average values of the post are as follows:
All : 2.3 post/day
Individual : 1.09 posts/day
Group : 1.23 posts/day

A two-sample t-test on the number of posts per day for individual user vs group users yielded the following results:

Attribute	Individual	Groups	T-value	P-value
Mean	1.09	1.23	**1.29748**	**.0383819**
Std. dev	0.7	0.63		

Table 5.3 Shows engagement through posts for individual vs group. Results are found to be significant at $p < .05$.

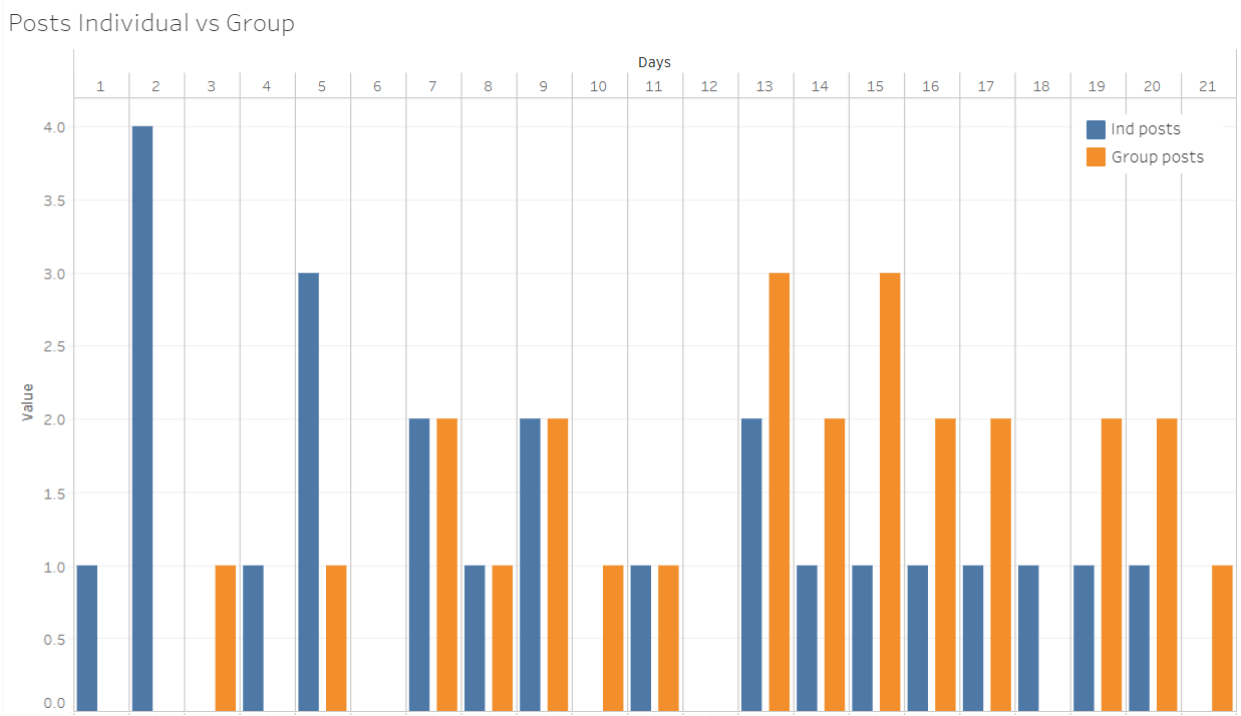


Fig 5.3 Shows comparison of posts made by individual members vs group members

We also plot the trend of posting frequency over the entire study duration. Individual members have higher post frequency during the initial period of the study. This decreased towards the end. Whereas the number posts made by group members are slightly increased as the time passed.

Furthermore, an ANOVA test is performed to analyze the effect of group size on post frequency

Attribute	Size 2	Size 3	Size 4	Size 5	F-value	P-value
Mean	1.833	1.166	0.62	0.6	2.1133	0.02143
Std. dev	0.41	0.381	0.274	0.24		

Table 5.4 Shows engagement through posts for different group sizes. Results are found to be significant at $p < .05$.

5.2.2 Number of Notification Conversions

Notifications are sent to an associated users' mobile device whenever there's an update to any activity in the system. The total number of notifications sent to the application during the study period were 4033, where individual members received 1654 notifications and group members received 2379 notifications. Average values of notifications per day are as follows:

All : 192 Notifications/day
Individual : 78 Notifications/day
Group : 113 Notifications/day

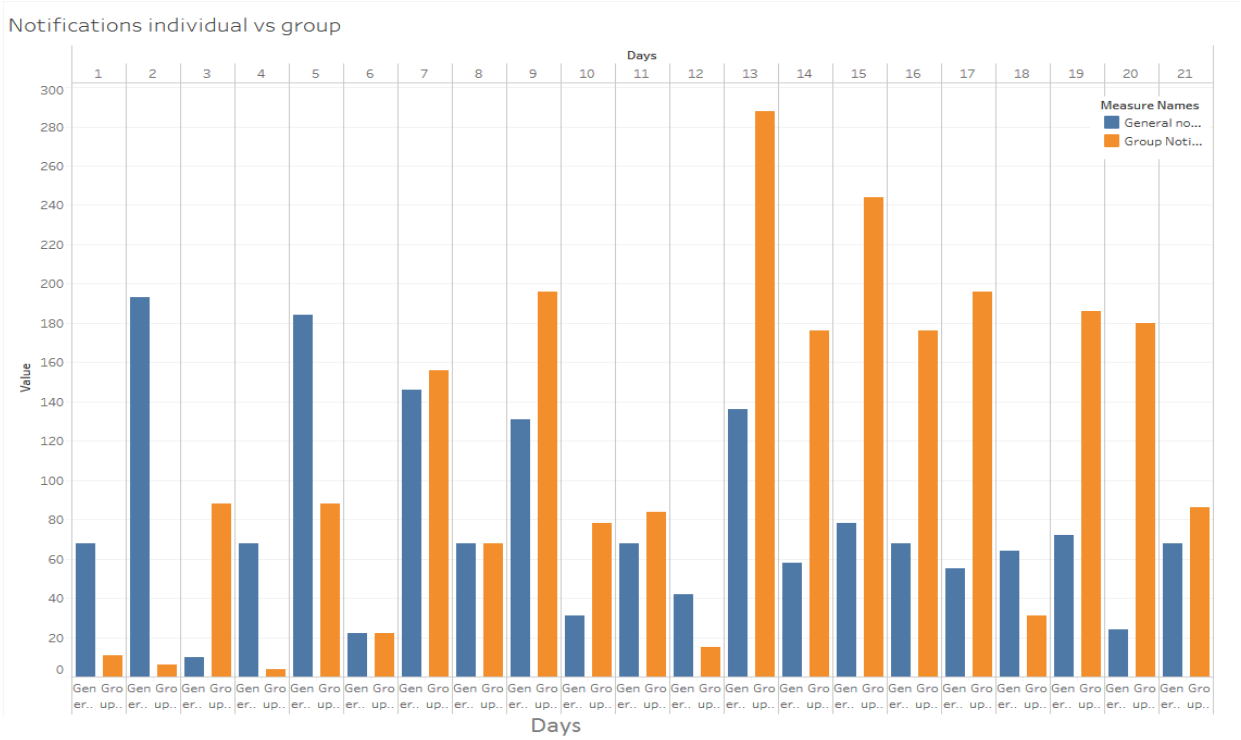


Fig 5.4 Shows comparison of notifications sent to individual user vs group users

The total number of notifications sent to group members is 43.8 % more compared to total notifications sent to individual members. Fig 5.4 and 5.5 shows the distributions of notifications of groups with different size over the entire study duration.

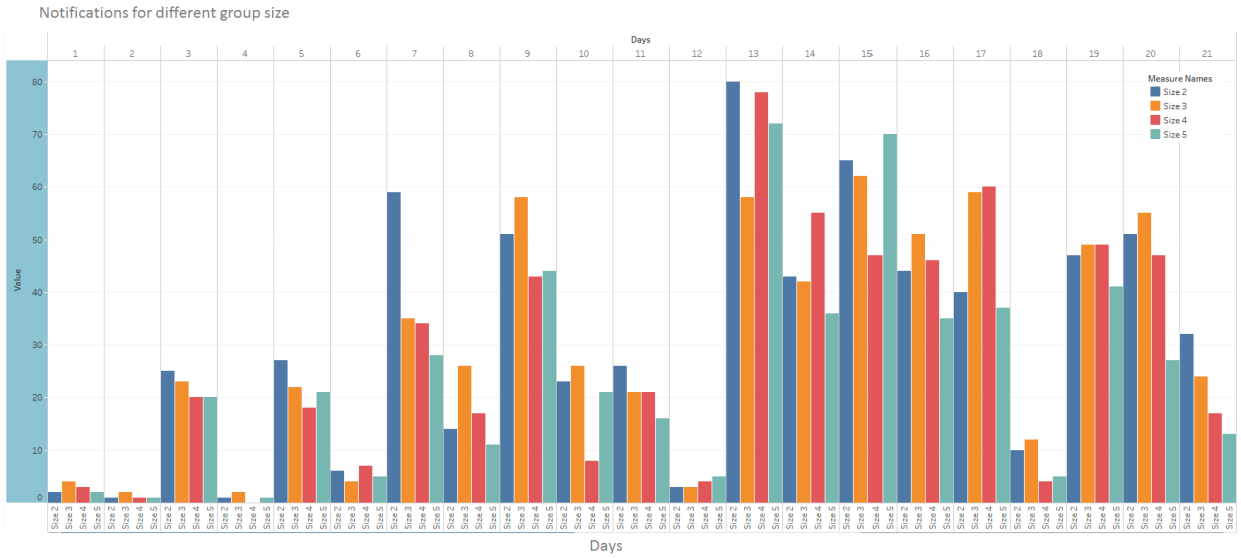


Fig 5.5 Shows comparison of notifications based on different group sizes

After receiving the notifications, how users react to them is an important indicator of engagement. Converting notifications to click is a fairly difficult task of user engagement. Users tend to dismiss a notification if they don't find it useful at first glance. In our proposed system, we tried to optimize the notifications based on previous learnings and couple them with group features. Below are the figures that show the conversion rates of each type of notification. Based on a two-sample t-test, the difference of conversions is statistically significant

Type	General notifications	Group notifications
Total sent	1654	2379
Converted to clicks	112	193
Mean	5.333	9.19
Std dev	3.26087	6.25794
% Conversion	6.7%	8.11%
P-value	0.007853	

Table 5.7: Compares conversion rate for notifications and t-test

5.2.3. Number of Comments

Comments can be added in Snapeve for every post. There are two types of post, 1. Generic to everyone 2. Targeted at group members. We analyzed the usage pattern of the comments and tested the significance to compare individuals and groups and compare group sizes. The following plot shows the comparison between generic comments and group focused comments.

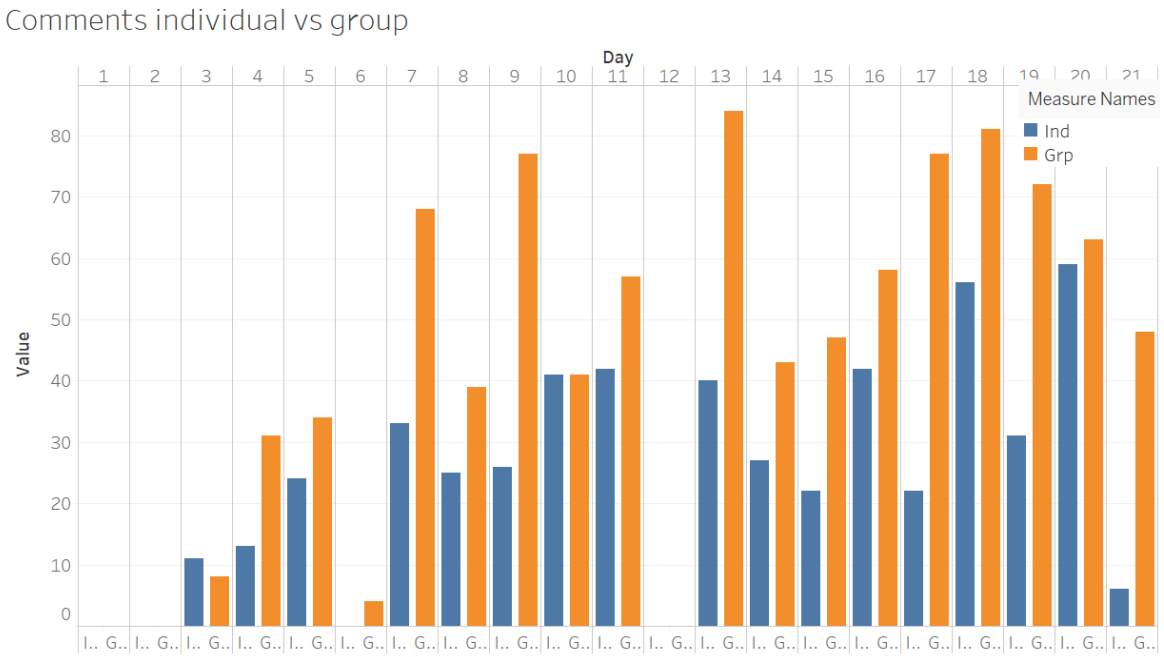


Fig 5.6 Shows comparison of comments for individual and group members

t-test on the above data gives the following results,

Attribute	Individual	Groups	T-value	P-value
Mean	31.470	38.041	**3.0182**	**.0041**
Std. dev	6.134	8.99		

Table 5.8 Shows engagement through comments for individual vs group

Further extending the analysis to the group size, using ANOVA, we get the following results

Attribute	Size 2	Size 3	Size 4	Size 5	F-value	P-value
Mean	41.91	36.6	29.5	27.17	**2.9782**	**.047636**
Std. dev	8.501	5.82	8.2	7.3		

Table 5.9 Shows engagement through comments for various group sizes, the result is significant with $p < .05$

5.2.4. Number of Invites

Invites are sent to other members in the app. Invites are sent to 1. anyone using the application, 2. group members only. There were total 261 invites sent during the study period, 123 sent by individual users and 138 by group members. The generic invites(default) that are sent from one user to any other user are plotted below along with invites sent by group members. The following plot shows the visualization of invites

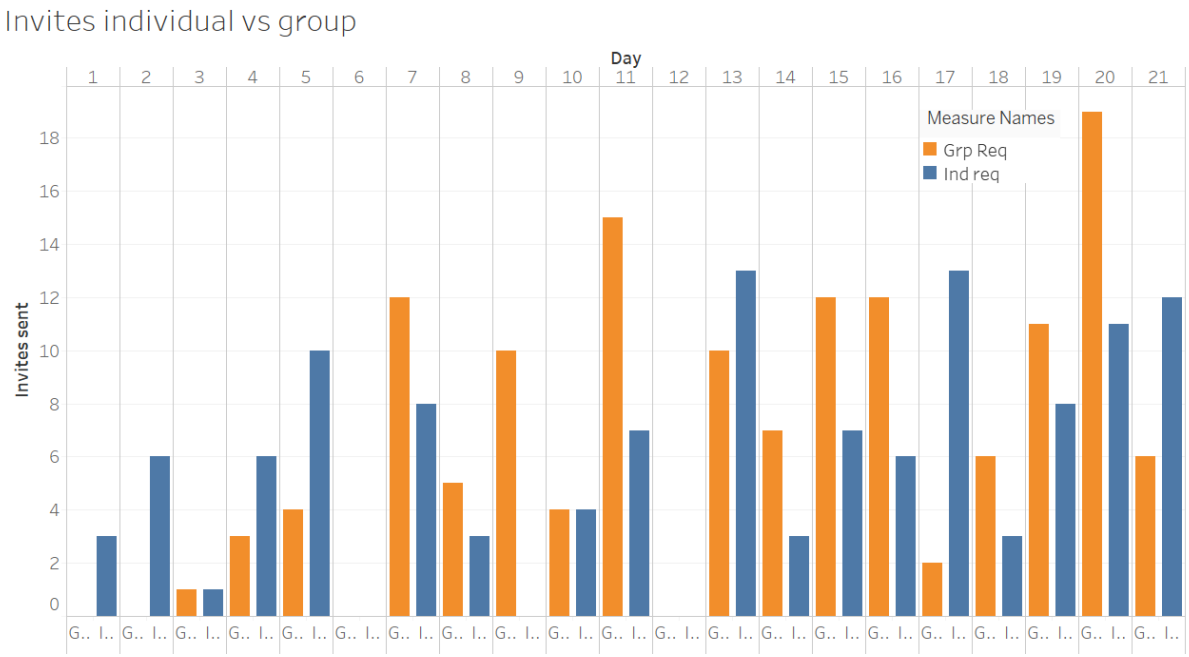


Fig 5.7 Shows comparison of invites for individuals and group

t-test performed on above data gives us the following results:

Attribute	Individual	Groups	T-value	P-value
Mean	5.9047	6.619	**1.9715**	**.0490**
Std. dev	4.229	5.5269		

Table 5.10 Shows engagement through invites for group vs individuals. The results are found to be significant at $p < .05$.

ANOVA test based on the group size gives us the following results

Attribute	Size 2	Size 3	Size 4	Size 5	F-value	P-value
Mean	5.52	5.49	5.13	5.60	**2.1572**	**.05737**
Std. dev	1.15	2.23	2.4	1.3		

Table 5.11 Shows engagement through invites for various group sizes

The results are NOT found to be significant at $p < .05$.

5.2.5 Number of group feature visits

Here we will observe the user behavior relating to group features based on the data logs collected by the backend of the application.

A. Profile page visits

There were total 1699 visits to the Profile page in the application during the study period. 714 belonged to individual "*Profile page"* by individual users and the remaining 985 were for "*Group profile page*" visited by group members. The figure below shows the distribution of the visits according to user type.

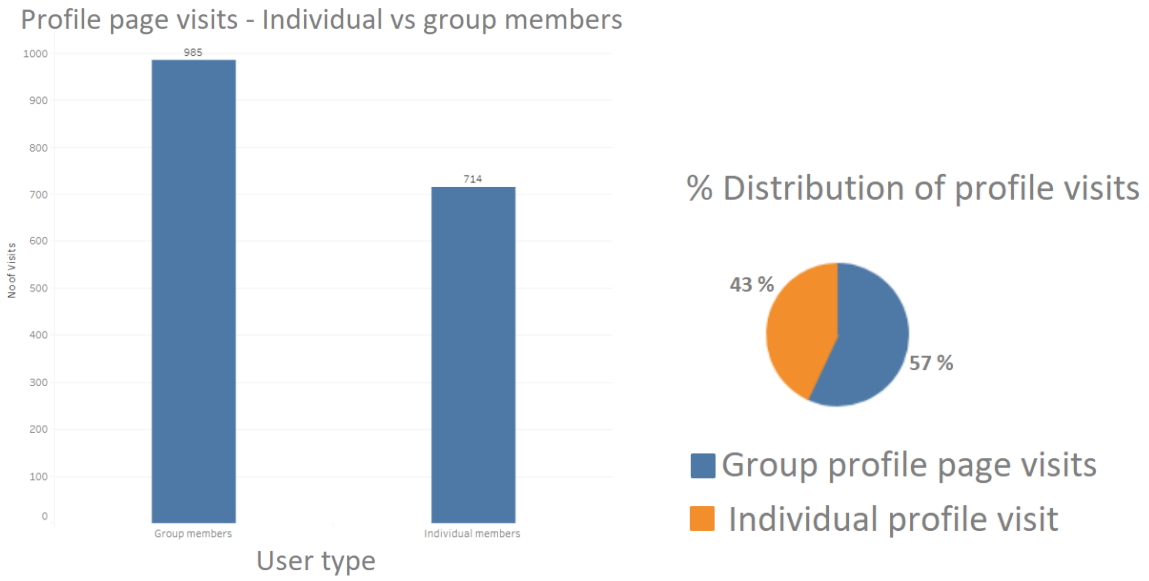


Fig 5.8: Shows the total number of profile page visits according to user type (Left) and percent distribution (Right).

The above distribution shows us the variation in visits to the profile page for different user type. The group profile page received 214 more visits during the period of study. Which can be considered as a significant difference given the size of the sample population. In addition, a two-sample t-test on profile visits in group vs individual setting gives us the following results:

Attribute	Individual	Groups	T-value	P-value
Mean	34	39.4	**0.39748**	**.031311**
Std. dev	9	6.1		

Table 5.12 Shows engagement through profile page visits for group vs individuals, the results are found to be significant at $p < .05$.

Moreover, we divided the group profile page visits into 4 categories depending upon the group size, which is shown in the pie chart below

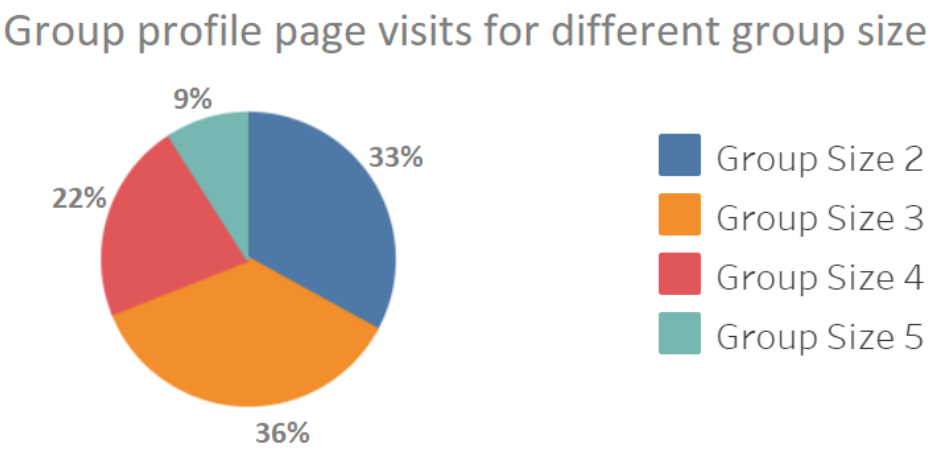


Fig 5.9: Shows the distribution of Group profile page visits according to group size

The pie chart shows the visits for group size 2 and 3 are dominant over the other two sized groups, and their share covers almost ⅔ of the total visits by all group members. ANOVA test on the following data gives us the following f and p values

Attribute	Size 2	Size 3	Size 4	Size 5	F-value	P-value
Mean	31.33	41.6	36.2	21.4	**1.3949**	**.027822**
Std. dev	8.3	11.8	8.7	4.2		

Table 5.13 Shows engagement through group profile page visits for various group sizes. The result is significant at $p < .05$

B. Leaderboard visits

The Leaderboard displays the current score of a user and/or current score of the group if the user is associated with a group. Usually, users visit the *leaderboard* page after they post an event or if they perform any action which gives them virtual rewards. The figures below show the distribution of visits to the leaderboard page according to user type followed by different group sizes.

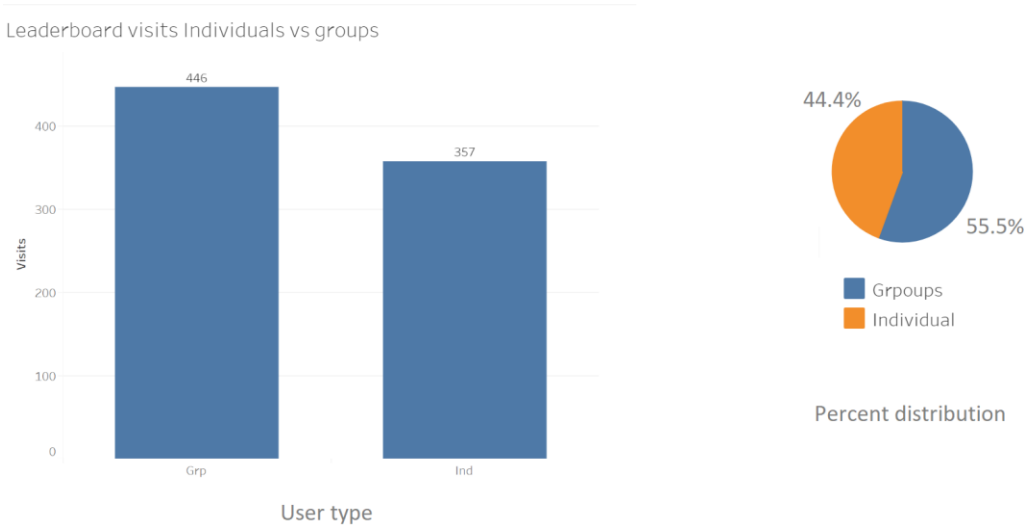


Fig 5.10: Shows the distribution of total leaderboard page visits according to user type (Left) and percent distribution (Right)

t-test results for the above data are as follows:

Attribute	Individual	Groups	T-value	P-value
Mean	17	17.84	**0.31448**	**.037351**
Std. dev	6.05	8.1		

Table 5.14 Shows engagement through Leaderboard visits for group vs individuals

The result is significant at $p < .05$.

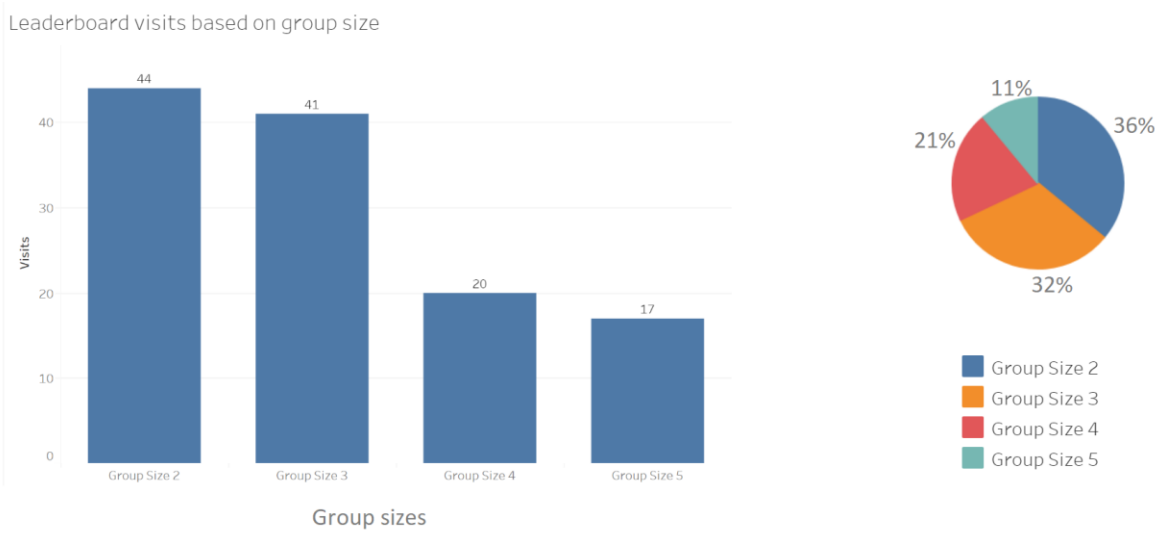


Fig 5.11: (Left)Shows the average number of visits by a user to leaderboard page from all group sizes, (Right) Pie chart shows the group-wise distribution of total visits

ANOVA test on the following data yields the following results:

Attribute	Size 2	Size 3	Size 4	Size 5	F-value	P-value
Mean	7.33	6.83	2.5	3.4	**1.4149**	**.029721**
Std. dev	2.8	3.1	0.8	1.2		

Table 5.15 Shows engagement through the Leaderboard visits for different group sizes. The result is significant at $p < .05$.

It can be observed that the number of visits to *the Leaderboard* page has similar statistics to that of profile page visits. i.e. group members show more frequency of visiting the leaderboard page as compared to individual members. Moreover, observing the frequency of group members according to different group sizes, it seems smaller groups are much more interested in check the leaderboard compared to larger groups.

C. Weekly rewards page visits

Weekly rewards are usually accessed during the weekends. The *weekly rewards* scheduler, in the backend of Snapeve, is scheduled to push notifications regarding the highest contributor every Friday at 7 PM local time. The frequency of visiting weekly rewards page is expected to be low compared to *the Profile page* or *Leaderboard* page since it is a weekly task and is often not check during the weekdays. Hence, the frequency of visiting this page is not considered in this research, rather the time spent by users on this feature is considered as a metric. Out of all the participants, 82% participants have checked the weekly rewards page all the times whenever rewards were declared at the weekend. Hence, only those participants' data is considered for this analysis, as other members who left the study early and could not access

this page. Since it was an analysis on duration rather than frequency, complete data is expected for all possible samples. Figures below further elaborate this:

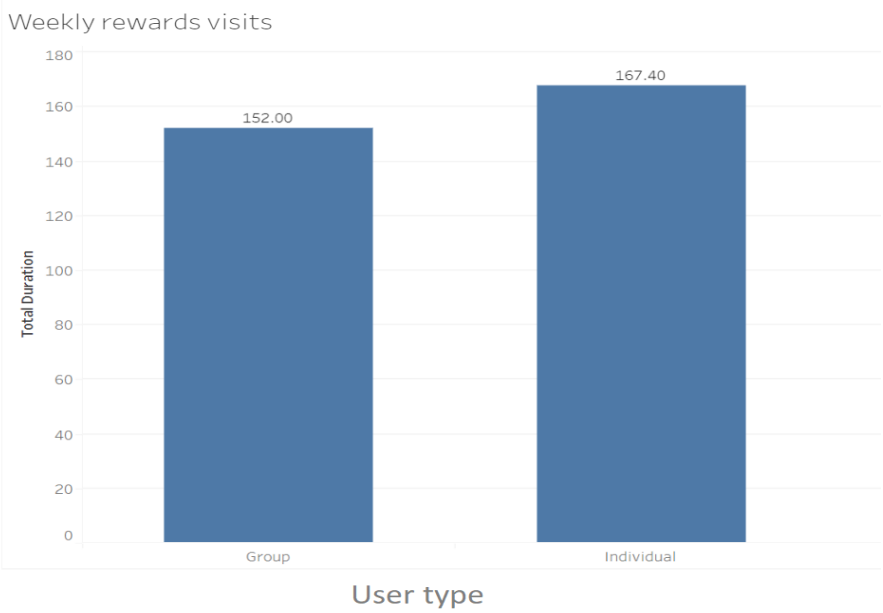


Fig 5.12: Shows the combined time spent on weekly rewards page according to user type

From the above figure, we can observe that there is a slight difference in the total time spent by users on weekly rewards page in both the user settings.

Attribute	Individual	Groups	T-value	P-value
Mean	7.23	6.68	**0.41142**	**.049315**
Std. dev	2.5	1.7		

Table 5.16 Shows engagement through Weekly rewards visits for group vs individuals

Although the t-test result is significant at $p < .05$, it is on the edge of the test analysis scale. This might be due to the possibility that the frequency of visit to weekly rewards page is low and most of the population have not visited the page during the period of study. The next figures explained the average times in different group settings.

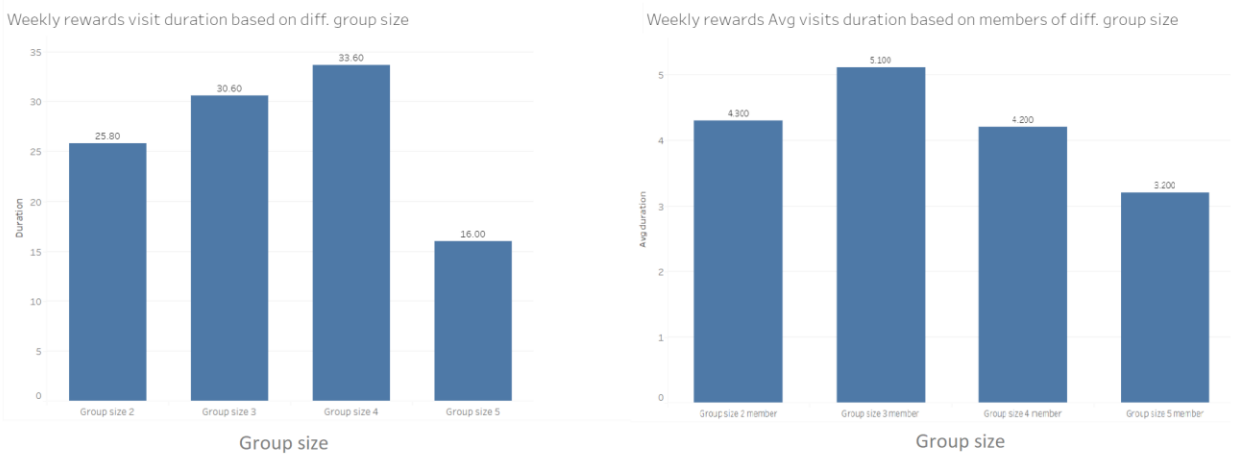


Fig 5.13: (Left)Shows the combined time spent on weekly rewards page by a group, and (Right)average time spent by a member of that group size.

The (left) figure shows that the group with size 4 has the highest visit rates with size 5 group the lowest and other group size lying in between the range. But when the calculation is done for an individual person from each group setting (Right), based on the number of participants in the group, it comes out that members in group size 3 spent more time compared to members of other group sizes, size 5 group being the lowest and members of other group size ranging in between. This tells us that, from an individual's perspective, group size 3 members have the longest time spent checking the weekly rewards.

5.3 Usability evaluation

An online survey was performed with the users of the system. To improve response rate, we pushed notifications to users' mobile devices and had the survey already embedded into the application of ease of access. Multiple notifications were pushed (1 in 2 days) to those participants who did not answer the survey to give them a reminder. As a result, out of the total of 46 participants, 35 participants responded, among them 17 individual users and 18 group members. The demographic information of the participants is mentioned below:

Age group	No of participants
18-24 years old	11
25-34 years old	13
35-44 years old	8
45-54 years old	1
55-64 years old	1
65+	0
Prefer not to disclose	0

Table 5.17: Age category - Online survey

Gender	No of participants
Male	23
Female	12
Other	0
Prefer not to disclose	0

Table 5.18: Gender category - Online survey

Android app use frequency	No of participants
Never	0
A few times a year	0
A few times a month	0
A few times a day	35

Table 5.19: Android app usage category - Online survey

The survey was designed using an online survey tool called *SurveyMonkey* with features of converting quantitative results into statistical data. The open-ended questions were listed and condensed considering factors such as frequency and persistence of the problem. The Likert scale questionnaire responses are mentioned below in table including usefulness and ease of use questions. The Likert scale ranges were from 1 to 5, 1 being least effective or strongly disagree and 5 being most effective or strongly agree. Responses to the survey from participants from both setting were collected and measured.

Q1: I think that I would like to use Snapeve frequently.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	1 2 9 8 15

Q2: I found Snapeve unnecessarily complex.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	27 8 0 0 0

Q3: I thought Snapeve was easy to use.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	1 1 5 19 10

Q4: I think that I would need the support of a technical person to be able to use Snapeve.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	34 1 0 0 0

Q5: I found the various functions on Snapeve were well integrated.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	0 2 8 15 8

Q6: I thought there was too much inconsistency in Snapeve	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	8 14 10 1 0

Q7: I would imagine that most people would learn to use Snapeve very quickly.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	2 3 8 12 10

Q8: I found Snapeve very cumbersome to use.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	24 4 5 1 1

Q9: I felt very confident using Snapeve.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	1 3 8 28 5

Q10: I needed to learn a lot of things before I could get going with Snapeve?	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	34 1 0 0 0

Q11: I would recommend Snapeve to a friend.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	0 0 1 22 12

*Q2, Q4, and Q10 are the reverse type of question where 1 means positive answer and 5 means negative, hence the values are normalized while calculating the following results.

Responses to each question above are measured and the t-test values are listed below.

Q. no	Individual mean	Group mean	Std dev (Ind)	Std dev (Grp)	T-value	P-value
Q1	4.76	3.79	1.20	1.33	3.1154	0.0199
Q2	4.12	4.67	0.10	0.18	0.911	0.0196
Q3	3.16	3.26	1.12	1.31	3.1822	0.0082
Q4	4.02	4.88	0.3	0.1	0.0010	0.0112
Q5	2.1	2.92	0.41	0.62	3.7423	0.0031
Q6	2.99	2.62	1.31	1.42	2.3122	0.0470
Q7	3.82	4.21	1.12	1.67	3.7353	0.0028
Q8	3.38	4.01	0.34	0.62	1.4421	0.0493
Q9	1.741	3.29	1.12	1.03	2.9153	0.0123
Q10	4.02	4.88	0.3	0.1	0.0050	0.0136
Q11	4.47	3.62	0.61	0.33	2.6405	0.0119
Avg	3.507364	3.8318	--	--	2.182027	0.018082

Table 5.20 Shows engagement t-test results for general usability survey questions

Below are the responses to the questions that were exclusive to group members

Q12: The group dashboard was effective (Scale 1-5)	No of participants
Strongly Disagree Disagree Neutral Agree	0 0 8 9

Strongly Agree	1

Q13: The leaderboard is effective in conveying the status of different individuals and groups.	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	0 0 1 10 7

Q14: The application interface is effectively in supporting interactions with your group members?	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	0 0 0 15 3

Q15: The system notifications are effective in informing you about your group members' activities?	No of participants
Strongly Disagree Disagree Neutral Agree Strongly Agree	0 0 1 15 2

To test the effect of group size on a user's satisfaction of group features, we employ the ANOVA test.

Q. no	Size 2 mean & std dev	Size 3 mean & std dev	Size 4 mean & std dev	Size 5 mean & std dev	F-value	P-value
Q12	3.96	4.37	3.33	3.23	0.4622	0.04575
	0.77	0.43	1.10	0.33		
Q13	4.12	4.47	3.88	3.71	0.8211	0.03257
	1.32	1.12	1.02	0.81		

Q14	4.82	4.25	4.92	4.41	0.9963	0.04683
	1.01	0.98	0.45	0.34		
Q15	4.60	4.81	3.9	3.69	1.0216	0.02338
	1.13	1.48	1.0	0.89		
AVG	4.375	4.475	4.007	3.76	0.8253	0.03713

Table 5.21 Shows engagement ANOVA results for group feature usability questions

In addition to the above questions, one descriptive question was included in the questionnaire for both the type of users. It was an open-ended question which enabled users to give their opinion apart from the questions based on Likert scales. Responses to this question are mentioned below:

Q16: Do you have any comments about Snapeve?

Responses from individual users	Responses from group users
● It could be better in messaging ● I want to be part of groups ● Excellent ● Good app ● It's nice	● It's great ● Nice app ● Useful ● Could have better rewards ● Needs better messaging in-app ● No way to communicate with emoticons ● No private chat ● Need dedicated group chat

Table 5.22: Summary of responses from users on the question *"Do you have any comments about Snapeve?"*

There were 14 users who did not provide specific comments.

From the above survey responses, we calculated the averages for responses which answers usability, satisfaction, and ease of use for the application. We found that people who are part of a group have positive averages indicating more satisfaction by being part of a group.

We found that users found the application easy to use and users found the application well integrated and simple. This result supports our hypothesis **H4a** stating that being part of a group impacts user satisfaction levels in a positive way.

In addition, we took the averages for survey responses comparing different group sizes and we found that people belonging to smaller groups have better satisfaction levels compared to people from larger groups. This also supports hypothesis **H4b:** group size affects user satisfaction levels in a positive way.

5.5 Summary

In summary, based on the t-test and ANOVA tests, all our main hypotheses were supported by the results from our study. Specifically,

Hypotheses	**Null Hypothesis(H_0)**
H1: User engagement improves when users are associated with a natural group **(Supported)**	User engagement levels are the same regardless whether the user is a part of a natural group or not.
H2: User retention rate improves when users are associated with a natural group **(Supported)**	User retention rates are the same regardless whether the user is a part of a natural group or not.
H3a: Group size affects user retention rate **(Supported)**	User retention rates are the same regardless of the sizes of their associated natural groups.
H3b: Group size affects user engagement **(Supported)**	User engagement levels are the same regardless of the sizes of their associated natural groups.
H4a Whether a user belongs to a natural group or not affects his/her overall satisfaction levels **(Supported)**	User satisfaction levels are the same regardless whether he/she belongs to a natural group or not.
H4b: The size of a group impacts its members' satisfaction levels **(Supported)**	User satisfaction levels are the same regardless of the sizes of their associated natural groups.

6. CONCLUSION

We conducted a research to study the impact of natural groups on community-based information sharing. As a part of the research, we custom built a mobile app called Snapeve which allow college students to share campus events in real time. We also conducted a 3-week user study to test several main hypotheses on the effects of natural groups on user retention, engagement and satisfaction. Our results indicate that users find more satisfaction, better engagement and longer retention if they are a part of a natural group as compared to an individual user using the same application. We also found that smaller groups performed better than larger groups in terms of their engagement and satisfaction in using the application. There is a wealth of information and research on collaborative information in organization level, but research on whether to enroll members of a natural groups together to improve retention and engagement is still new. The study results can help us understand a new way to elicit user interests and encourage user participation in community-based sharing information.

7. FUTURE WORK

Understanding how voluntary information sharing can be studied further in a group environment. Factors like demography, the nature of a group, the structure of a group and familiarity of a group may play an important role, In the future, we can extend our study to take these factors into consideration.

Our current mobile application is Android-based. This has prevented us from recruiting a larger number of participants. In the future, we will develop an iPhone app for both campus event sharing and community-based solution for fighting opioid crisis.

8. REFERENCES

Savolainen, R. (2017). Information sharing and knowledge sharing as communicative activities. Information Research, 22(3), paper 767. Retrieved from http://InformationR.net/ir/22-3/paper767.html (Archived by WebCite at http://www.webcitation.org/6tTRz0IcS)

Danny Yang Chang Ho. 2011. Towards a Framework for Collaborative Information Exchange: Exploring Effects of Time Pressure and Task Difficulty on Group Cognition Behaviors and Implications for Computer Supported Collaboration in Healthcare. Ph.D. Dissertation. University of Maryland at Baltimore County, Catonsville, MD, USA. Advisor(s) Anthony F. Norcio. AAI3459930

L. Chuang and S. Chiu, "Exploring collaborative information exchange in social media," *2016 International Conference on Applied System Innovation (ICASI)*, Okinawa, 2016, pp.1-4.doi: 10.1109/ICASI.2016.7539819.URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7539819&isnumber=7539562

L. Chuang and S. Chiu, "Encouraging Knowledge Sharing Among Green Fashion Communities," *2017 International Conference on Green Informatics (ICGI)*, Fuzhou, 2017, pp.141144.doi:10.1109/ICGI.2017.43URL:http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8117095&isnumber=8117062

S. Sifi and R. Alouane, "Esprit-social network: An internal collaborative platform for the students of ESPRIT," *2016 IEEE Global Engineering Education Conference (EDUCON)*, Abu_Dhabi,2016,pp.563-567.doi:10.1109/EDUCON.2016.7474607,URL:http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7474607&isnumber=7474513

Widen, G., & Hansen, P. (2012, December 4). Managing collaborative information sharing: Bridging research on information culture and collaborative information behaviour. Retrieved from http://www.informationr.net/ir/17-4/paper538.html #.W-h0b5NKg2x

Tanja Aitamurto, Hélène Landemore & Jorge Saldivar Galli (2017) Unmasking the crowd: participants' motivation factors, expectations, and profile in a crowdsourced law reform, Information, Communication & Society, 20:8, 1239-1260, DOI: 10.1080/1369118X.2016.1228993

Anne Ekholm, Issues of supporting communities in mobile contexts, SIGGROUP Bulletin, December 2012. Retrieved from https://dl.acm.org/citation.cfm?doid=990017.990024

Hossain, Mokter. (2012). Users' motivation to participate in online crowdsourcing platforms. ICIMTR 2012 - 2012 International Conference on Innovation, Management and Technology Research. 10.1109/ICIMTR.2012.6236409.

Lim, Yujin & Ahn, Sanghyun & Kang, Sanggil. (2012). Mobile group system for social networking services on smart campus. 251-252. 10.1145/2401603.2401658.

X. Zhang *et al*., "Incentives for Mobile Crowd Sensing: A Survey," in *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 54-67, Firstquarter 2016. doi: 10.1109/COMST.2015.2415528 url: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7065282&isnumber=7393921

Anesthesiol, K. J. 2015, "T test as a parametric statistic". doi:10.3897/bdj .4.e7720.figure2f https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4667138/

Restor Dent Endod, 2014, "Analysis of variance (ANOVA) comparing means of more than two groups" https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3916511/

Steiner, I. D. 1966. Models for inferring relationships between group size and potential group productivity. Behavioral Science 11:273–283

Hill, G. W. 1982. Group Versus Individual Performance: Are N + 1 Heads Better Than One? Psychological Bulletin 91:517– 539.

Kittur, A., and Kraut, R. E. 2008. Harnessing the wisdom of crowds in wikipedia: quality through coordination. In Proceedings of the Conference on Computer Supported Cooperative Work, 37–46. ACM Press.

Tausczik, Yla & Wang, Ping & Choi, Joohee. (2017). Which Size Matters? Effects of Crowd Size on Solution Quality in Big Data Q&A Communities.

Vassilev, I., Rowsell, A., Pope, C., Kennedy, A., O'Cathain, A., Salisbury, C., & Rogers, A. (2015). Assessing the implementability of telehealth interventions for self-management support: A realist review. *Implementation Science,10*(1). doi:10.1186/s13012-015-0238-9

How to Calculate and Improve Customer Retention Rate. (n.d.). Retrieved from https://www.salesforce.com/hub/service/customer-retention-rate/The Landmark

Jarman, S. (2017, January 08). The Must-Follow Best Practices for Your Push Notifications. Retrieved from https://hackernoon.com/the-must-follow-best-practices-for -your-push-notifications-5f878565d2a9

DOSTÁL, O. (n.d.). Better UX Through Microinteractions. Retrieved from https://www.toptal.com/designers/product-design/microinteractions-better-ux

Toporek, A. (2017, August 7). How to Calculate Customer Retention Rate. Retrieved from https://customersthatstick.com/blog/customer-loyalty/how-to-calculate- customer-retention-rate/

Ramshaw, A. (2018, August 31). How to Analyze Survey Data in Excel. Retrieved from https://www.genroe.com/blog/analyze-survey-data-in-excel/11483

Get these best practices on how to analyze your survey data. (n.d.). Retrieved from https://qualaroo.com/marketers-guide-surveys/analyze-survey/

Leahy, J. (2004). Using excel for analyzing survey questionnaire. Retrieved from https://learningstore.uwex.edu/Assets/pdfs/G3658-14.pdf

Babich, N. (2018, February 12). A Comprehensive Guide to Mobile App Design. Retrieved from https://www.smashingmagazine.com/2018/02/comprehensive-guide-to-mobile-app-design/