The impact of using a discussion support system on discussions

A Bachelor's graduation thesis.

Luc Schoot Uiterkamp – University of Twente

Supervised by:

Prof. Dr. F. Van der Velde

Dr. X. J. Dai

Table of contents

Abstract

Smart technologies that are made to assist people in their daily lives are becoming more abundant due to their decrease in price an increase in capabilities. The Living Smart Campus initiative develops projects that use smart technology and tests them in the small ecosystem of the University of Twente campus. The discussion support system under review in this thesis is an app made to assist participants of discussions in collaborative decision making processes. Sixteen participants spread over four groups have tested the app by discussing twice; once with app and once without. A questionnaire was filled in by every participant after each discussion, with eleven questions that each represented a different aspect of the discussion. An indication of an improvement of the focus on the topic at hand by the use of the app was found. An indication of an effect of the discussion topic on the discussion effectiveness and depth of the discussion was found, as well as a significant effect of the topic on the clarity of communication flow. A structural equation model was made that represents how the questionnaire items relate to the use of the app and the discussion topic. An indication of an effect of the use of the app and the topic on group related items was found and should be investigated in a study set up to identify differences between group related and individual items.

1. Introduction

Smart social technologies play an ever increasing role in our lives. With the decrease in size and price and increase in performance and capabilities of smart electronics, their incorporation into society becomes more widespread and more deeply interwoven. Smart social systems are designed to implicitly or explicitly assist or serve us. More than 25 years ago, Mark Weiser (1991) predicted such ubiquitous systems in his paper 'The computer for the 21st century'. He thought of smart contextually aware devices that could assist us in an implicit way. Today, these systems are already present and quite advanced. For example smartphones try to present data, reminders and notifications based on what it thinks suits the user at the time. It can postpone notifications until a certain meeting is over or increase its volume if it senses it is in a noisy environment. Some smart technologies are aimed to serve a broad interest, for example smart sensors in roads that can sense not only the traffic driving on the road but also the state of the road (Miller, 2016). This currently serves as a maintenance indicator but could also warn other traffic directly. Other systems are designed to serve more specific purposes, such as supporting participants of a discussion in decision making processes. Many different support systems that fulfil this purpose exist, and one of them is the subject of this thesis. The road sensors and the discussion support app were developed by the Living Smart Campus project on the University of Twente, an initiative that develops smart technologies and tests them on the University of Twente campus.

1.1 Living smart campus

Innovations such as the road sensors and the discussion support system naturally need to be developed and tested before they are used in society. The Living Smart Campus (LSC) project is an effort to test subtly present technologies on the University of Twente campus before they are used in the real world. The devices under test are generally in an early stage of development, on the cutting edge of technology and are aimed at making the campus smarter and at unobtrusively making a whole array of processes run smoother. The small ecosystem present on the campus enables the technologies to be tested in a way that is somewhat representable to the ecosystem the solution will be applied to eventually. Additionally, using the University campus as testing ground provides a controllable environment while providing valid results because the campus acts like a small real-world society.

The project started as an initiative of the Centre for Telematics and Information Technology (CTIT) at the University of Twente but has grown out to be a multidisciplinary network of projects, all implementing technical solutions to mostly social questions. As a subproject of the Living Smart Campus initiative, an application made to improve decision making in discussions is tested. This thesis will test the perceived improvement of discussions when using the application. Research into the app

is done in collaboration with other researchers, who look into different aspects of the subject. Other researchers are doing an objective analysis on the labelling within the app and are investigating the effect the app has on turn taking in discussions, the number of words used per comment and the coherence of tagging. The aim is to draw a conclusion regarding the effectiveness of the app in improving discussions using the combined data of the subjective data collected in this study and the objective data collected by other researchers.

1.2 Discussion app

The decision support system covered in this thesis is a mobile application aimed at improving discussions by simplifying collaborative decision making processes and providing a better overview of the discussion in hindsight. The app records the discussion for each participant individually, which means each participant of the discussion uses their own mobile device with the app loaded. During the discussion, the participants can indicate that an argument, a counterargument, a proposition or an off topic comment is made on the app. Because each participant makes these indications individually, they may differ between participants. The indications are marked on the recording for later reviewing. By documenting the nature of the comments, the structure of the meeting becomes clearer, the grounding for the decisions made during the discussion is documented in a more structural way and the overview of processes after the meeting is better. This improves documentation and enables comparison of proposals within the meeting with their respective counterproposals. Off topic comments are identified, and can be discarded or treated with lower priority.

The app was initially developed by research group Tech-cico at the Université de Technologie de Troyes in France, under the name MMProject (Xinghang, Nada, & Guillaume, 2015). There are IOS and Android versions of the app available, but the Android version is no longer maintained and therefore out of date.

1.3 Decision making

This thesis, and more generally the subproject of the LSC project that this thesis is part of, is focussed on the improvement of collaborative decision making processes. Decisions can be made alone or in groups, and in numerous ways. Different studies have shown that collaborations are more efficient and effective than individual efforts in decision making and people in groups “generally learn faster, make fewer errors, recall information better, make better decisions, and produce a higher-quality product than do individuals” (Johnson & Johnson, 2014, p. 262). Groups outperform individuals by such an extent because members fill in each other’s gaps in knowledge and provide new insights by combining knowledge from different sources. This process of interaction is known as process gain, and yields more options to choose from and provides better solutions than one single person might have

thought of. Individuals may be wrong about a certain aspect, which is recognised and filtered out by group members, improving the information quality (Johnson & Johnson, 2014; Owen, 2015).

There are different styles of decision making in groups. A decision can be made by a group leader after hearing opinions of different group members, by averaging the members' opinions, by majority voting and by collaboratively deciding on a solution that the whole group endorses. The scale of working together to achieve results ranges from no collaboration to being entirely built upon collaboration, and both strategies have advantages and disadvantages. An advantage of little collaboration is that it takes very little time, but there is rarely any support for the decision and hardly any of the advantages of group discussion are utilised. Averaging participants' contributions or casting a majority vote are decision making styles that provide room for every individual to contribute to the discussion. The influence each participant has on the outcome is large and advantages of group discussions are utilised. However, averaging and voting result in averages of propositions made during the discussion, which are rarely creative or innovative. Collaborative decision making is focussed on unanimous backing rather than compromising as is the case with voting or averaging. The unanimous backing results in universal ownership and agreement regarding the solution, which improves the willingness to implement it. This makes the choice made more valuable than the individual proposals (Johnson & Johnson, 2014; Owen, 2015).

A disadvantage of collaborative decision making is the time investment that is necessary. The decisions made collaboratively are by nature backed by the whole group, but coming to a conclusion endorsed by each member takes time. If the process of coming to a solution is not efficient, the increase in time costs might not be worth the increase in decision quality and if it is, a reduction in time while maintaining the quality of the decision will be of great benefit to the group.

1.4 Decision support systems

A decision support system can be used to make the process of reaching consensus more efficient and easy. A decision support system (DSS) is "a computer-based system, typically interactive, that is intended to support people's normal decision-making activities" as defined by Yates, Veinott, and Patalano (2003, p. 39). This is a fairly broad and vague definition, to encompass the numerous definitions for decision support systems that are developed for specific contexts and target groups. There exist a great number of decision support systems that are made to fulfil very specific tasks in very narrowly defined parameters (see Razmak and Aouni (2015) for a number of examples), but the discussion app does not fit those criteria. It has a broader application and is aimed at supporting the process of getting to a decision rather than proving support for a specific decision. This falls under the category 'group decision support systems'.

Traditionally, group decision support systems focus on group behaviour and group processes rather than on the content of the discussion. They are made to avoid groupthink, to resolve conflicts effectively and provide tools to make processes such as voting easier. In general, they provide behavioural patterns to follow in order to reach a grounded decision (Poole, Holmes, Watson, & DeSanctis, 1993; Yates et al., 2003). For more information on groupthink, conflicts that may occur and support tools for working together see Johnson and Johnson (2014). Shim et al. formulate group support systems somewhat more loosely and more in line with the support system app used. They define them as "systems [to] (…) enhance communication related activities of team members engaged in (…) cooperative work" (2002, p. 117). They define time, space and level of group support as characterising factors of support systems. The support system used in this study is synchronous in time, meaning that the participants of the discussion will take part in the discussion at the same time rather than one after another. They are also in the same space, enabling face-to-face contact. The level of group support is not well defined for the current support system, it offers support to individual participants as well as the whole group.

The support system app presents the user with a model to use but does little more, the term 'discussion support system' might therefore be more appropriate than decision support system. To use generally accepted terms, it is a model- and communication driven group support system (Baumeister & Striffler, 2015; Shim et al., 2002) and fits under the 'general' decision support system label used in a group context, as defined by Yates et al. (2003).

The simplicity of the currently studied support system may lack the effectiveness more elaborate systems can offer, such as step by step guides to conflict resolution or the prevention of inhibiting factors such as groupthink or freeloading. However, the simplicity of the system also makes it easy to use and therefore requires little attention to successfully use. More elaborate decision support systems can detract from decision quality because they distract the focus from the discussion. The decision support system competes for attention with the discussion at hand and the amount of attention required to use the decision support system may be significant. This shift of attention from the discussion to the support system causes the discussion to lose depth and focus which deteriorates the decision making process and results in a poorer quality discussion with a poorer quality decision. However, as said, the decision support system used in this study is relatively simple and is not expected to show extraneous cognitive load to a great extent.

1.5 Current study

The discussion support app is an example of a minimally obtrusive computer that provides assistance to a social situation. As mentioned above, the LSC project is aimed at improving the quality and ease of life on the campus by providing assistance in such social situations. Because of the academic nature of the university campus, a lot of meetings take place, whether from project groups of students, teaching staff or researchers. This makes it a very fitting environment to test a discussion support system, while also providing an impact on the efficiency of all kinds of meetings. For these reasons, the LSC project incorporated the discussion support system to further their development.

The above-mentioned improvements the app can make are generally focussed on analysis after the meeting is over. The current study investigates whether the act of structuring the discussion makes the discussion itself more effective. Four groups of four participants had discussions, once with and once without app, each half an hour. Afterwards, they filled in a survey in which they scored several aspects of the discussion the app is thought to influence. This survey data constitutes the subjective data. Using transcriptions of the discussions, objective data was collected.

The study consists of two parts, an exploratory and corroborative part. The former is aimed at exploring in what ways the discussion is influenced by the use of the discussion support app, according to the participants of the discussion. For this part, the subjective survey data will be used. The latter part of the study is aimed at comparing the results of the subjective exploration with objective measures taken from the discussions. Factors that have been significantly impacted by the use of the app will be looked into more deeply and will be compared to objective data. This comparison can provide objective support to the claims based on subjective data.

2. Factors influenced by the use of the app

The factors that are thought to be influenced by the use of the app are listed and elaborated upon here. There are eleven factors in total, each paragraph in this section elaborates on one or more of these.

2.1 Amount of contributions (factor 1)

The amount of positive contributions each participant brings to the table naturally has an impact on the quality of the discussion. If more people weigh in more, the discussion is likely to be better (Johnson & Johnson, 2014). Communication oriented support systems tend to be focussed on optimising contribution of all participants in a discussion by promoting active participation and proving guidelines on turn taking and contributions (Poole et al., 1993; Yates et al., 2003). The app used in this

study does not feature explicit guidelines but rather a communicational model to follow. Participants of a discussion tend to adapt their discussion style based on the decision support system if one is available because this requires less effort than fitting the discussion support system to match the discussion style (Todd & Benbasat, 1991, 1992). In other words, if a discussion support system is used, it is easier to adapt the discussion to the constraints and freedom offered by the support system than to try to work around them to make the support system fit the discussion. For this study, this means that participants should be inclined to base their comments on the available labels: argument, counterargument, proposition and off topic comment. Basing a contribution on one of the labels should, with the exception of 'off topic' comments, yield useful additions to the meeting. Because the participants are expected to use the provided labels as guideline for their discussion, more of these useful contributions to the discussion are expected to be found when the app is used. To determine if the amount of arguments, counterarguments and propositions increases with the use of the app, the perceived amounts are measured. It is expected that the amount of arguments, counterarguments and propositions will be larger when the app is used then when it is not used because of the participants' expected inclination to use the provided structure.

2.2 Focus (factor 2)

Keeping the discussed subject related to the general topic of the discussion while still being able to discuss details is a fine balance but one that must be kept because an off topic discussion cannot be productive. Usually discussion support systems rely on users to keep track of the topic, the systems merely provide tools but cannot assess sematic information such as relevance of the topic. In the discussion support app, users are required to actively indicate off topic comments. This leads to quick identification of the off topic comments. Comments are labelled individually by each participant, so there might not be a consensus on whether a certain comment is off topic. However, if a comment is off topic it can be expected to be identified as such by a majority of the group. In most productive meetings, participants will not be inclined to continue on an off topic subject voluntarily and if a majority thinks a certain comment off topic, they are expected to steer away from this topic. Therefore, the identification an off topic comment by a majority of the group can help keep the meeting focussed on the topic at hand.

As mentioned before, the focus of a discussion may be negatively influenced by the presence of a decision support system. However, this effect is expected to be quite small because of the simple nature of the app used in this study. Therefore, the positive effect that the off topic labelling has is expected to create an overall increase in focus on the topic during discussions.

2.3 Content clarity (factor 3)

Not being aware of the relevant advantages, disadvantages and opinions makes it more difficult to assess a standpoint or proposal (Yates et al., 2003). Specialised discussion support systems can provide background information on certain topics to improve the knowledge about a certain topic but there is more to understanding the content of the discussion than knowing much about the topic at hand. An important part of an effective discussion is knowing what is talked about at that moment (Johnson & Johnson, 2014). The app cannot show this but by labelling the content of the discussion, the app raises awareness on important comments regarding the current subject. As mentioned, the labelling is done per participant and may therefore be different between participants. However, the awareness of the importance of a certain comment is heightened, even it is not labelled correctly by everyone because the participants that have labelled it correctly are aware of its importance and will continue on that subject. Additionally, a wrong label may also heighten the awareness of the importance of the comment. Owen (2015) state that a problem in collaborative decision making is that not everyone is aware of their own and each other's conceptual frames. They propose that individuals pay more attention to what is right about someone else's proposal than to what is wrong, to increase awareness of overlap in the respective conceptual frames. The app facilitates this by raising awareness to arguments that vouch for a certain proposal. Apart from facilitating collaborative decision making, being aware of important statements and relevant arguments clarifies what others have said about the subject and makes it easier to assess and formulate arguments, counterarguments and proposals (Johnson & Johnson, 2014). For these reasons, content clarity is measured in this study and is expected to be greater when using the app than when not using the app because of the heightened awareness of important remarks.

2.4 Decision quality (factor 4)

The main point of any decision support system is to somehow support people in the making of a decision. Lehto and Nah describe this as "play[ing] a useful role in reducing biases and otherwise improving decision quality" (2006, p. 230). In many decision support systems this goal may be achieved by defining parameters within which a certain decision should lie or by presenting a possible solution based on a mathematical model. Communication oriented support system such as the app under review in this study may achieve this goal in a more indirect manner. The decision quality is expected to improve because the more effective communication enables a higher quality decision. Additionally, factors mentioned above are generally parts of 'improving decision quality'. For these reasons, perceived decision quality is measured in this study. It is expected to be better when using the app than when not using the app.

2.5 Clarity of communication flow (factor 5)

As mentioned, effective discussions require effective communication. The clarity of communication flow, the overview of who is talking to whom, plays a role in this. Group communication takes place through a number channels, each channel being a communication line between two individuals. In this study, groups of four individuals participated in the discussions, yielding twelve possible communication channels. Figure 1 shows a diagram of the possible communication channels in a discussion with four participants such as in this study. Within these channels, the communication may be one-to-one, one-to-many or even many-to-one and clarity and overview of these channels is essential to keep the discussion on track (Johnson & Johnson, 2014). Support systems, especially those made for asynchronous discussions, assist in clarifying the communication flow by clearly labelling who has said what. Also synchronous but spatially separated discussions such as video meetings benefit from this.

The app has no automatic system to indicate who is commenting or to whom the remark is addressed. However, because each participant labels the current comment type, they are being made aware of the current state of the discussion, who is talking and who is addressed. This increases the overview of the communication flow in general, and therefore the clarity of the communication flow is thought to be better when using the app than when not using the app.

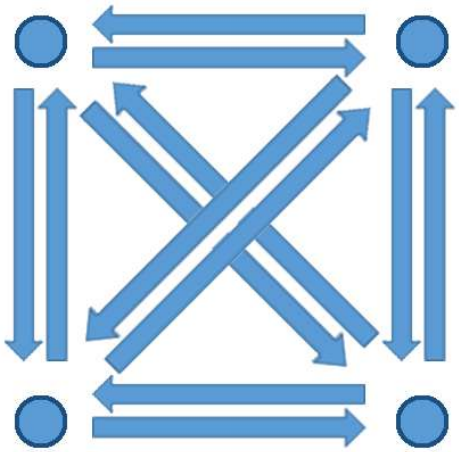

Figure 1. The possible communication channels for a group of four people.

2.6 Structure (factor 6)

Structure in a discussion is an important part of its effectiveness. Structured discussions have a systematic procedure for coming to a conclusion. This helps prevent forgetting important steps leading to that conclusion and ensures a correct order of steps in arriving there. Support systems may provide this structure directly or may provide procedural instructions that help participants of a discussion achieve structure (Lehto & Nah, 2006). Group decision support systems may also give structure to a discussion by implicitly providing a structure to adhere to, as is the case in the app. Participants of a discussion tend to adapt their discussion style to the support system because it is easier to adapt to the support system than fit the system to the discussion (Todd & Benbasat, 1991, 1992). In practice, this means that the discussion using the support system should have a structure resembling the support system structure. For this study, this means a discussion consisting of proposals, arguments and counterarguments. Because the app implicitly brings structure to the discussion, the amount of structure is expected to be greater when using the app than when not using the app.

2.7 Participation (factor 7) and collaboration (factor 8)

Participation and collaboration are essential factors in an effective collaborative discussion and its outcome (Johnson & Johnson, 2014; Owen, 2015; Yates et al., 2003). Current support systems attempt to improve participation and collaboration by equalising the participation across participants and by promoting participation and collaboration (Reinig & Mejias, 2014). According to Kolfschoten, Lukosc, and Leimeister (2013, p. 290), collaboration may be supported by “setting agenda’s or procedures, proposing techniques, moderating discussion[s], guarding quality, analyzing and summarizing results, creating awareness” and they propose that “ways of monitoring and predicting collaborative activities, ways to assess and determine interventions and ways of adapting collaboration environments to guide participants in the collaboration process and to alter collaborative behavior” are the next steps in support systems with a focus on collaboration.

As the app is not solely focussed on collaboration and is meant to be generally applicable, the suggested features for the future are not present in the current app. The app does present a discussion model that promotes collaboration because of the ‘proposal-argument-counterargument’ structure. Additionally, the increased clarity of content and communication flow should make it easier to participate and collaborate (Yates et al., 2003). For these reasons, and because the participation and collaboration are fundamental to the discussion and the app’s influence on it, both participation and collaboration are measured as their own measure. They are both expected to be greater when using the app than when not using the app.

2.8 Depth (factor 9), quality (factor 10) and effectiveness (factor 11) of the discussion

One of the main goals of the app is to make discussions more effective, which means the task of discussing several topics and reaching a conclusion with consensus from each individual were performed with sufficient quality without taking up too much time. Support systems usually achieve this indirectly by providing as much information and guidelines as possible so that users can spent more time on the matter at hand. The depth of the discussion plays an important role in the quality and effectiveness of the discussion, because if a certain topic is discussed on a deeper level, more information is available for that topic and a well-grounded decision can be made. Because of the increased clarity of content and increased amount of structure during the discussion, it is expected that the discussion depth is greater when using the app than when not using the app. Quality and effectiveness consist of more than merely depth however, so in addition to depth of the discussion, they will be measured separately as well and expected to be greater when the app is used than when the app is not used.

3. Overview research questions and hypotheses

The eleven factors listed above are expected to be influenced by the use of the discussion support app during the discussion. The first part of the study focusses on the perceived improvement of these factors made by the app during discussions. This will be measured by administering a questionnaire after both a discussion with and without the support app. The second part of the study is a follow up analysis on the results of the first part in which the subjective measures will be compared to objective measures gathered from recordings of the discussions to investigate if the objective measures corroborate the findings from the first part of the study. This corroboration is tested both in direct comparison and in a structural model.

Each hypothesis for the first part of the study is mentioned in the relevant paragraph above and reviewed in order here:

1. The amount of arguments, counterarguments and propositions are larger when the app is used then when it is not used**.**
2. Focus on the topic during discussions is greater when using the app than when not using the app.
3. Content clarity is greater when using the app than when not using the app.
4. Discussion quality is better when using the app than when not using the app.

5. Clarity of the communication flow is better when using the app than when not using the app.
6. The amount of structure is greater when using the app than when not using the app.
7. Participation is greater when using the app than when not using the app.
8. Collaboration is greater when using the app than when not using the app.
9. The discussion depth is greater when using the app than when not using the app.
10. Decision quality is better when using the app than when not using the app.
11. Discussion effectiveness is greater when using the app than when not using the app.

4. Methods

4.1 Design

The app was tested with four groups of four participants, where each group had two discussions; one where they used the app and one where they did not use the app. There were two predefined topics for the discussions: 'Design a bicycle parking system for the university' and 'Discuss a website for student and employee communication'. In the discussions where participants used the app, each participant had their own tablet on which the app was loaded. The use of the app in the discussions and the topics of the discussions were balanced. Each combination of app use and topic was used exactly once in a first discussion and once in a second discussion. Table 1 shows the 2 x 2 conditions and table 2 shows the exact distribution of app use and subjects in during the data collection.

Table 1

The four experimental conditions used in this study.

		With app	Without app
Discussion subject	*Website*	*Website with app*	*Website without app*
	Bicycle stands	*Bicycle stands with app*	*Bicycle stands without app*

Table 2

Balanced distribution of app use and discussion subjects between first and second discussion.

	1st discussion		2nd discussion	
Group:	*App:*	*Subject:*	*App:*	*Subject:*
1	With app	Bicycle stands	Without app	Website
2	Without app	Bicycle stands	With app	Website
3	With app	Website	Without app	Bicycle stands
4	Without app	Website	With app	Bicycle stands

4.2 Participants

To achieve the distribution shown in table 2 with four participants per group, 16 participants were needed. Participants were recruited via the human subjects management system SONA. Students were rewarded three SONA credits to participate in the study. A total of 20 participants signed up to take part in the study, 19 of which actually took part in it. The first group, where one participant was missing, was used as a pilot test. There were 16 participants that took part in the actual data collection. Of these sixteen participants, 13 were female and 3 were male. Three groups had one male in them and one group had exclusively females. The average age was 20 *(range: 18-22)*. All participants were psychology students at the University of Twente.

4.3.1 Materials for part 1 of the study

Pen and paper questionnaires were handed out after each discussion. They were labled with 'discussion 1' or 'discussion 2' to distinct the first and the second time the group met and with 'with app' or 'without app' to indicate if the app was used during that discussion. The questionnaire was developed based on the eleven factors described in section 2 of this paper that are thought to be influenced by the app. One question was made for each factor and all questions were 5 point Likert scales. The full questionnaire can be found in appendix A.

The iPad and Android version of the app differ slightly in interface and functionality. To keep the collection of the data consistent, only one version of the app was used. Because the iPad version of the app was more up to date, only this version was used. The iPads used were iPad 2's and iPad Airs. A microphone headset plugged into each iPad was used to record sound for each participant.

In the app, three indicators that serve as guidelines for subtopics to be covered during the discussion are shown on the left side. The four buttons labelled ‘Argument’, ‘Counter-argument’, ‘Proposal’, and ‘Off topic’ are shown on the right side.

Figure 2 shows a screenshot of the app with three stages of the discussion on the left and the four buttons that can be pushed by participants on the right.

For the study performed by the co-researchers, two cameras were set up during all the discussions in such a way that each camera had two participants’ faces clearly in shot as well as most of the other two participants so that it can be clearly determined to whom a participants is talking. Furthermore, a dictaphone was used to record the discussions as well as the interviews performed for the co-researchers’ study.

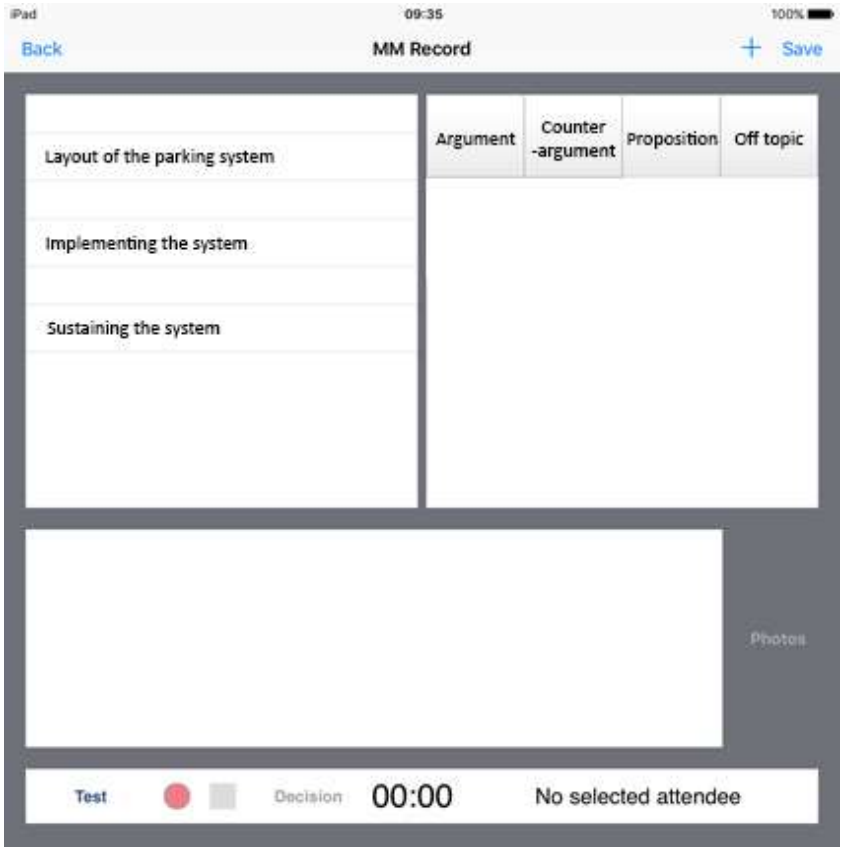

Figure 2. A screenshot of the support system app. Three guidelines for the discussion are shown on the left, the four buttons indicating ‘Argument’, ‘Counterargument’, ‘Proposition’ and ‘Off topic’ are shown on the right.

4.3.2 Materials for part 2 of the study

Using the objective data collected by the other researchers working on this project, an objective measure for focus was made. An increase in useful contributions and a decrease in off topic comments is expected if the focus on the topic at hand increases. Therefore, the objective measure for focus during the discussions is defined as the mean number of propositions, arguments and counterarguments minus the number of off topic comments. In mathematical terms this is:

$$Combined\ objective\ measure\ for\ Focus = (Propositions + Arguments + Counterarguments\ /\ 3) - Off\ topic\ comments$$

where each of the terms indicate the number of comments of that type made in each discussion. In the second part of this study, the propositions, arguments, counterarguments and the combined objective measure for focus are compared with the subjective focus. The combined measure is added to the existing structural equation model to identify how the objective measure for focus contributes to the model.

4.4 Procedure

Before the experiment, participants were told there would be two meetings and that they would have a discussion both meetings. Participants were told to discuss a predefined topic, either 'Design a bicycle parking system for the university' or 'Discuss a website for student and employee communication'. The topic was never twice the same for one group and both topics are used an equal amount of times in the first and second discussion as well as in the with and without app conditions. See table 2 for the exact distribution. For both discussions, the participants were told to read the instructions. These instructions can be found in appendix B. Preceding the discussion where the app was used, the participants were instructed to label arguments, counterarguments, propositions and off topic comments when they felt like they were made by pressing the relevant button. They were told that, for the tagging, it did not matter who made the remark; they were instructed to press the button regardless of who was speaking. The button lay-out can be seen in figure 2. When all participants were done reading the instructions, a researcher indicated the start of the discussion by clapping their hands. This clap was later used to sychronize the audio streams of the cameras, iPads and dictaphone. The participants had exactly 30 minutes to discuss their subject, and, after 25 minutes, were told they had five minutes left. When the 30 minutes were over, a researcher told the participants to stop discussing. One participant at the time filled in the questionnaire and was interviewed for a co-researcher's study. During the interview and the filling in of the questionnaire, the other participants were asked to wait in a separate room. After each participant had filled in the questionnaire and taken the interview, the participants were told they may leave.

4.5 Data analysis

The data will be tested for normality using the descriptive statistics. The answers to the questionnaire will be compared between the condition that used the app and the condition that did not with a paired samples test. A non-parametric test will be used to compare the questionnaire items between conditions because the paired samples t-test is not robust to deviations from normality on small sample sizes. De Winter and Dodou (2010) have shown that Likert scales can be analysed with both a paired samples t-test and a Wilcoxon test with similar reliability. To test for confounding variables, the topic of the discussion and the order of the discussion will also be tested.

A factor analysis will be used to determine how many aspects of the discussion are measured by the questionnaire. If there is more than one, the items will be sorted and assigned to their appropriate factors. The factors, their respective items, and the use of the app will be modelled in a structural equation model to determine the impact of the app use on the perceived quality of the discussion.

In the second part of the study, factors that have appeared to have been significantly impacted by the use of the app will be compared to objective measures taken from the discussion transcripts as described above. The objective and subjective measures will be compared directly by correlation analysis and indirectly by the additional value the objective values may contribute to the structural model.

5. Results of part 1

Descriptive statistics indicate a normal distribution of the data with no abnormally large skewness or kurtosis values for any of the measured items except for the 'focus' item, which has a relatively high positive kurtosis value, indicating a narrower and peakier distribution than normal. A full table with descriptive statistics can be found in appendix B. Table 3 is a frequency table of the scores of the eleven questionnaire items, with and without the use of the app during the discussion. The frequencies are colour coded: darker shades indicate a higher frequency. The item numbers are the items in the order as they appear on the survey, which can be found in Appendix A.

Table 3

Colour coded frequency table comparing the 'without app' condition with the 'with app' condition. Darker colours represent higher frequencies.

		Condition											
		Without app						**With app**					
Score		*1*	*2*	*3*	*4*	*5*		*1*	*2*	*3*	*4*	*5*
Items	**1**	1	3	6	6	0	**1**	0	2	10	3	1
	2	0	3	7	6	0	**2**	0	1	5	10	0
	3	0	2	8	6	0	**3**	0	1	10	5	0
	4	0	1	8	7	0	**4**	0	1	7	8	0
	5	0	3	4	9	0	**5**	0	2	3	8	3
	6	0	6	8	2	0	**6**	0	6	4	5	1
	7	0	2	3	9	2	**7**	0	0	1	12	3
	8	0	0	0	11	5	**8**	0	0	1	9	6
	9	0	5	9	2	0	**9**	0	1	13	2	0
	10	1	2	7	6	0	**10**	0	2	9	4	1
	11	0	1	4	10	1	**11**	0	0	9	7	0
Total		2	28	64	74	8	***Total***	0	16	72	73	15

Note: items 1-11 are the items in order as they appear on the survey in Appendix A.

Table 4 shows results of a Wilcoxon signed rank sum test on the questionnaire items with the use of the app and the discussion topic. Only results that are close to being significant are shown. The focus on the topic during discussions where the app was used were not significantly better than during discussions where the app was not used but there is an indication that an effect is present. The discussion effectiveness and depth of the discussion showed a trend towards a significant improvement in discussions with the website topic over the discussions with the bicycle rack topic. The clarity of communication flow was significantly better during discussions with the website topic than during discussions with the bicycle rack topic. The questionnaire items showed no significant difference related to the order of the discussions.

Table 4

Wilcoxon signed rank sum test of the factors expected to be influenced with the use of app and the topic. Only results approaching significance are shown.

	Factors	**Z**	**P (two tailed)**
App	Focus with app – Focus without app	1,732[a]	0,083265
Topic	Discussion effectiveness website – Discussion effectiveness bikes	1,941[a]	0,052204
	Clarity of communication flow website – Clarity of communication flow bikes	2,126[a]	0,033471
	Depth website – Depth bikes	1,890[a]	0,058782

a. Based on positive ranks

A maximum likelihood factor analysis shows that the questionnaire measures two factors, which can be seen in table 5. Here, the eleven questionnaire items are sorted with their respective factors. Therefore, the two factors are incorporated as separately in the structural model, which can be seen as model 1 in figure 3.

Table 5

Maximum likelihood factor analysis matrix, rotated using Varimax rotation with Kaiser Normalization. Item load indicates the item load for each item, marked loadings indicate the appropriate factor for that item.

Items	**Item load**	
	Factor 1	Factor 2
Focus	,993	-,116
Clarity of communication flow	,537	-,070
Discussion quality	,534	-,175
Discussion effectiveness	,530	,027
Amount of structure	,334	-,214
Depth	,317	,161
Decision quality	,262	,040
Participation	,253	,967
Amount of contributions	,173	,846
Clarity of content	-,137	,270
Collaboration	-,140	,204

Note: Rotation converged in 3 iterations.

Figure 3 shows a structural equation model with the effects that the discussion topic and the use of the app have on the two factors found in the factor analysis. The two factors are predicted by the subjective measures as defined in table 5. Table 6 shows the path values, regression values, covariance values and interaction values for model 1 as shown in figure 3.

A chi square test on model fit showed the null hypotheses of a correct model should not be rejected (p = 0,440).

Figure 3. Structural equation model 1, showing the subjective measures predicting factor 1 and factor 2 as defined in table 5, the direct effects the topic and the use of the app have on the factors and the mediator effect the discussion topic has on the effect between the use of the app and the factors. The values associated with the paths can be found in table 6.

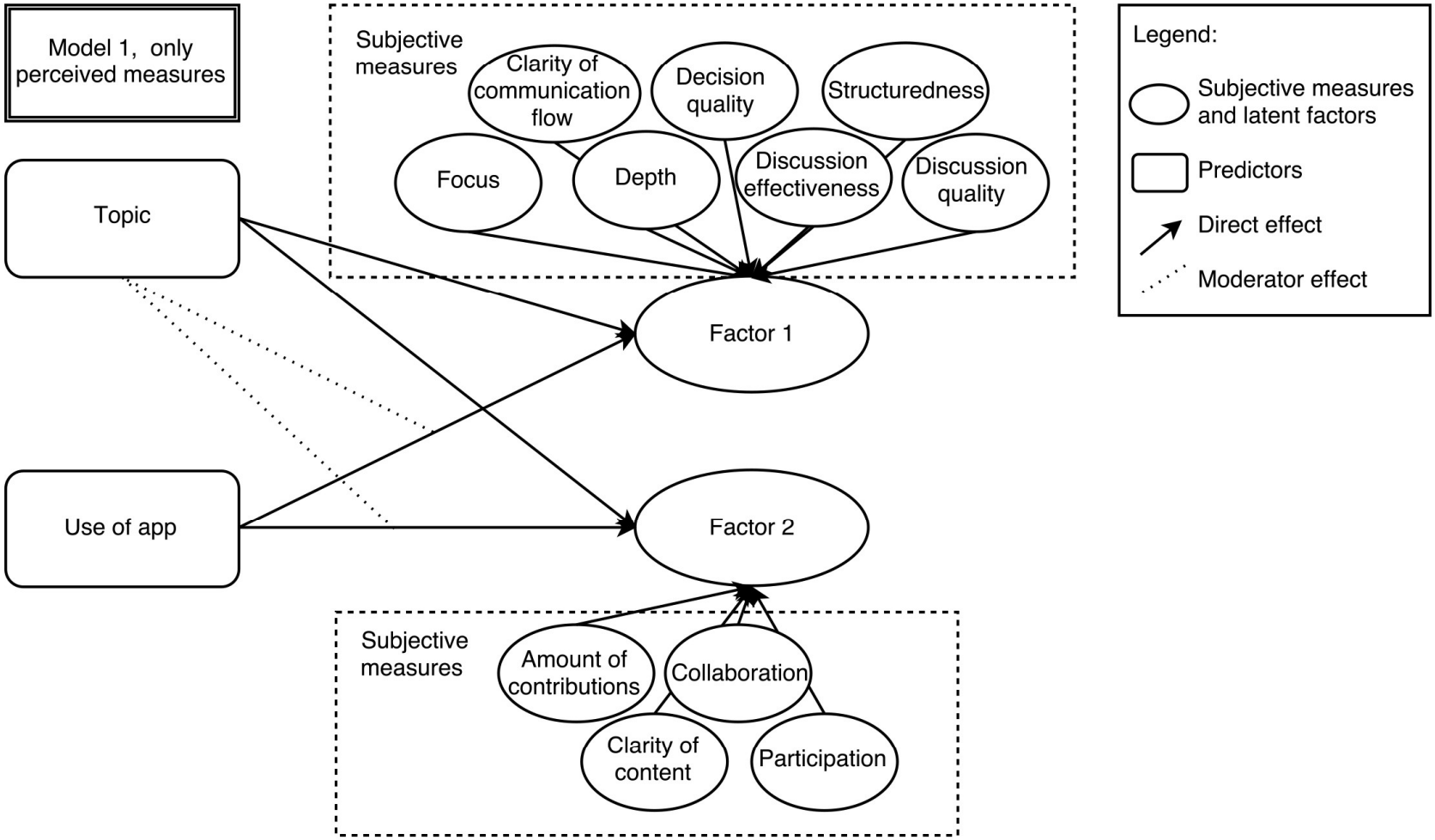

Table 6

Structural equation model path values, regression values and covariance tables for model 1 as shown in figure 3. The model paths indicate the path value that each item has in relation with its corresponding factor, the regressions indicate the regression analyses of the factors with the discussion topic and the use of the app, the covariance indicates the covariance between factor 1 and factor 2 and the interaction indicates the influence of the interaction effect of the discussion topic and the use of the app on factor 1 and factor 2.

Model paths

		Estimate	Std.Err	Z-value	P-value
Factor 1					
	Discussion quality	1,000			
	Discussion effectiveness	0,944	0,362	2,610	0,009
	Clarity of communication flow	1,344	0,531	2,531	0,011
	Amount of structure	0,994	0,491	2,027	0,043
	Focus	1,859	0,593	3,134	0,002
	Depth	0,433	0,320	1,353	0,176
Factor 2					
	How many contributions	1,000			
	Clarity of content	0,166	0,135	1,233	0,218
	Participation	1,241	0,526	2,358	0,018
	Collaboration	0,149	0,150	0,988	0,323

Model regressions

		Estimate	Std.Err	Z-value	P-value
Factor 1					
	Discussion topic	0,278	0,161	1,726	0,084
	Used app	0,217	0,130	1,666	0,096
Factor 2					
	Discussion topic	0,192	0,266	0,720	0,471
	Used app	0,101	0,229	0,440	0,660

Covariance

	Estimate	Std.Err	Z-value	P-value
Factor 1 ~~ Factor 2	0,018	0,042	0,422	0,673

Interaction

		Estimate	Std.Err	Z-value	P-value
Factor 1					
	Discussion topic x Used app	0,125	0,207	0,602	0,547
Factor 2					
	Discussion topic x Used app	0,549	0,433	1,267	0,205

Note: in this table 'Estimate' denotes the path estimates in the structural equation model (see figure 3), 'Std.Err' denotes its standard error, Z-value denotes the number of standard errors from the mean, and the P-value denotes the probability value for the statistic.

6. Discussion of the results of part 1

The results of the two parts of this study will be discussed separately. Results of the first part will be discussed here, the results and discussion of the second part as well as a general discussion will follow afterwards. Discussion of some of the general results will be omitted in this part and elaborated upon in the general discussion.

The focus item on the questionnaire had a relatively high kurtosis value but all tests used allow for this. For the paired samples test, the Wilcoxon signed rank test was used which allows for non-normal distributions. The structural equation modelling is not sensitive for slight deviations from normality and is therefore appropriate here.

In table 3, a decrease in frequencies of lower scores and increase in frequencies of higher scores can be seen in the 'with app' condition compared to the 'without app' condition. In general the scores appear higher on 'with app' condition. The Wilcoxon signed rank test shown in table 4 shows that the focus on the topic at hand during the discussions when the app was used might be higher than when it was not used. The increase in focus is not significant but does form an indication that such an effect might be present. Discussion order and discussion topic were subjected to the same test, which revealed a significant effect of the discussion topic on the clarity of communication flow and an indication of an effect of the discussion topic and the discussion effectiveness and depth of discussion. The discussion order had no significant impact on any items.

Table 6 shows that the use of the app and the topic both have an impact on factor 1 with a significance level that hints at a real effect. The effect found here is not significant, but is clearly a lot better than the effects on of both the use of the app and the topic on factor 2. No correlation was found between factor 1 and factor 2, which provides additional support for the distinction between the factors.

The goal of the second part of this study is to look into the subjective measures that are likely to be significantly affected by the use of the app during discussions. Only the focus on the topic at hand fulfils that criterion and will therefore be the only item subject to additional investigation in the second part of the study. The discussion topic had a suspected influence on a number of items, but will not be looked into further since the discussion topic falls outside the scope of this thesis.

7. Results of part 2

Table 7 shows Pearson correlations between the objective measures for focus and the subjective measure for focus. The subjective measure for focus is not significantly correlated with the combined objective measure for focus.

Table 7

Pearson correlations between the objective measures for focus and the subjective measure for focus.

	Correlations				
	1	**2**	**3**	**4**	**5**
1 Amount of proposals					
2 Amount of arguments	,449**				
3 Amount of counterarguments	-,097	,649**			
4 Amount of off topic comments	,151	,258	,227		
5 Objective focus combined *	,623**	,896**	,606**	-,024	
6 Subjective focus	,179	-,207	-,330	,076	-,174

*. For definition see discussion of results of part 1.
**. Correlation is significant at the 0.01 level (2-tailed).

Figure 4 shows structural equation model 2 with the effects the discussion topic and the use of the app have on the two factors found in the factor analysis. The two factors are predicted by the subjective measures as defined in table 5, and factor 1 is additionally predicted by the combined objective focus measure. Table 8 shows the path values, regression values, covariance values and interaction values for model 2 as shown in figure 4.

A chi square test on model fit showed the null hypotheses of a correct model should not be rejected ($p = 0,092$).

Figure 4. Structural equation model 2, showing the subjective measures and objective measure predicting factor 1 and factor 2 as defined in table 5, the direct effects the topic and the use of the app have on the factors and the mediator effect the discussion topic has on the effect between the use of the app and the factors. The values associated with the paths can be found in table 8.

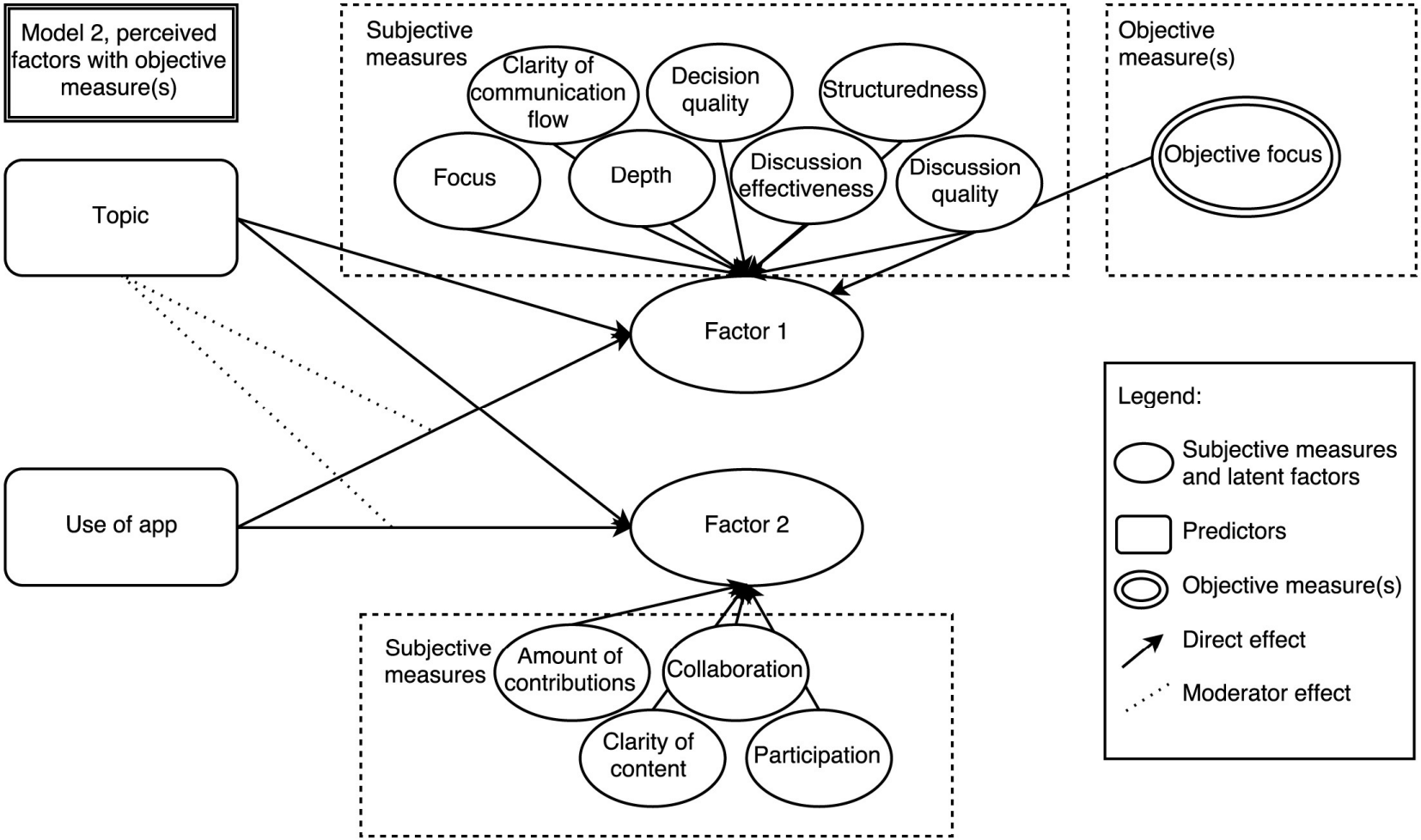

Table 8

Structural equation model path values, regression values and covariance tables for model 2 as shown in figure 4. The model paths indicate the path value that each item has in relation with its corresponding factor, the regressions indicate the regression analyses of the factors with the discussion topic and the use of the app, the covariance indicates the covariance between factor 1 and factor 2 and the interaction indicates the influence of the interaction effect of the discussion topic and the use of the app on factor 1 and factor 2.

Model paths

		Estimate	Std.Err	Z-value	P-value
Factor 1					
	Discussion quality	1,000			
	Discussion effectiveness	0,938	0,362	2,588	0,019
	Clarity of communication flow	1,366	0,533	2,561	0,010
	Amount of structure	1,003	0,500	2,006	0,045
	Focus	1,882	0,602	3,123	0,002
	Depth	0,425	0,323	1,314	0,189
	Objective focus	-2,446	2,685	-0,911	0,362
Factor 2					
	How many contributions	1,000			
	Clarity of content	0,167	0,134	1,244	0,213
	Participation	1,227	0,516	2,378	0,017
	Collaboration	0,148	0,151	0,981	0,326

Model regressions

		Estimate	Std.Err	Z-value	P-value
Factor 1					
	Discussion topic	0,269	0,159	1,698	0,089
	Used app	0,217	0,130	1,672	0,094
Factor 2					
	Discussion topic	0,196	0,268	0,731	0,465
	Used app	0,102	0,232	0,440	0,660

Covariance

	Estimate	Std.Err	Z-value	P-value
Factor 1 ~~ Factor 2	0,015	0,043	0,345	0,730

Interaction

		Estimate	Std.Err	Z-value	P-value
Factor 1					
	Discussion topic x Used app	0,125	0,207	0,602	0,547
Factor 2					
	Discussion topic x Used app	0,549	0,433	1,267	0,205

Note: in this table 'Estimate' denotes the path estimates in the structural equation model (see figure 3), 'Std.Err' denotes its standard error, Z-value denotes the number of standard errors from the mean, and the P-value denotes the probability value for the statistic.

8. Discussion of the results of part 2

None of the objective measures for contribution correlate significantly with the subjective measure for focus. Moreover, the combined objective measure for focus does not correlate with the subjective measure. The addition of the objective measure for focus did not have a significant impact on factor 1 and did not significantly change the model as a whole.

The lack of coherence between the subjective and objective measure for focus implies that the participants' view on the focus on the topic at hand is different from that of the researchers.

9. General discussion

9.1 Factor analysis

The factor analysis indicated that the subjective items loaded on two factors. The difference between these factors is not unambiguous but items loading on factor 1 tend to be group related items, whereas the items loading on factor 2 are more individual in nature. Examples of clear group oriented items in the first factor are discussion quality, discussion effectiveness, amount of structure, depth and decision quality. Even though the opinion of these matters may differ per individual, they are properties of the discussion as a whole, not an individual contribution. Only the clarity of communication flow seems an individual item. The items loading on the second factor are based on individual behaviour instead of the discussion as a whole. The participation, amount of contributions and clarity of content are all individual-centred items, they depend on how one participant behaved during the discussion. The collaboration however, seems a more group oriented item, relying on interaction between individuals. Despite the minor discrepancies in both factors, the general theme of factor 1 is considered 'group related' and factor 2 is considered 'individual'. They will therefore be named 'group related' and 'individual' accordingly. For consistency in this paper, they shall remain 'factor 1' and 'factor 2' in figures and tables.

The level of support was defined by Shim et al. (2002) as a defining factor for discussion support systems. The factor analysis in this study confirmed this distinction and shows that the app has two distinct effects, one on group based processes and one on individual processes.

9.2 The focus measure

The focus measure was predicted to increase with the use of the app because the app was expected to heighten the awareness of off topic comments and the importance of on topic comments. However, the objective measure for focus does not confirm an increase in on topic comments or decrease in off topic comments. The addition of the measure to the model did not change other values significantly and had no significant predicting power for factor 1. Furthermore, there is no correlation between any of the tags that were predicted to increase in frequency or the combined objective measure with the subjective focus measure. The subjective measure indicates that participants felt that the discussion was more focussed with the use of the app while the objective measure indicated that this was not the case. This may be the result of the chosen objective measure for focus. It was reasoned that in a more focussed discussion, less off topic comments and more on topic contribution would be made. However, it may be the case that participants made the same number of off topic comments but were made aware of them as predicted and switched to another topic. This would explain the lack of reduction of off topic comments as well as the suspected increase in subjective focus. Discussion support systems based on awareness of usefulness of contributions do not seem abundant, although there are decision making theories concerning awareness of the knowledge base in a team (Lehto & Nah, 2006). The awareness of comment types may be a foundation for a support system and could provide a useful addition to the already available tools.

9.3 Use of app and topic

Neither the use of the app nor the discussion topic had a significant impact on factor 1 or factor 2. However, the p-values of the effect the app and topic had on factor 1 (0,084 and 0,096 respectively) were an order of magnitude larger than the p-value of their effects on factor 2 (0,471 and 0,660 respectively). While both are still not significant, this observation is at least a reason to look into the differences between group related items and individual items in a study that is set up specifically for this task. Explanations for the observed indication of a possible effect between the use of the app and the discussion topic with factor 1 are described below to aid future research into this matter.

The use of the app and the topic may both have a bigger impact on the group aspects of a discussion and less on the individual aspects. The impact that the use of the app has on the group related factors can come from a general feeling of improvement of the discussion but not a specific sense of improvement in their own contributions and participation. The app was not aimed to provide either group support or individual support but was made to do both. However, only for the group related items an indication for an effect was found, which, if there is in fact an effect for group related processes, means this support system also has only one functional level of support.

Because this study used two substantially dissimilar topics for the two discussions, one topic might simply have been more suited for group discussions. The topic can have an influence on the effectiveness of a discussion support system and on the discussion itself. The participants might feel that the discussion with one topic went better than one with another while their participation and effort remained the same. This would explain the apparent influence of the topic on the group related items and its apparent lack thereof on the individual items.

A moderation effect of the discussion topic between the use of the app and the factors was expected because the discussion topic had a significant effect on a number of items in the questionnaire and because decision support systems in general can have varying effectivity between different types of discussion and discussion topics (Lehto & Nah, 2006). The interaction analysis in the structural equation model showed no significant moderation effect on the use of the app for either of the factors. It appears the topic had no effect on the effectivity of the app.

10. Possible improvements

If this study is to be repeated, care should be taken in deciding on recruitment methods and the amount of time each group has for their discussion. The participants in this study formed temporary groups and had very little or no experience as a group prior to the study. For this reason the groups were not mature, which impedes the group's capability to perform optimally (Johnson & Johnson, 2014). A setup more resilient for this type of error may include some or exclusively groups that are used to working together. Seeing as groups that are used to working together are potentially a large target group for the app, accounting for maturity will increase the validity and usefulness of the results.

Because reaching a unanimous backing can take a fairly long time, groups should be given ample time for their discussion. If participants feel like they have little time left, they might feel pressured to conclude their discussion without unanimous backing. Premature concluding of the discussion might lead to use of averaging or voting to make decisions and conflict avoiding techniques to avoid confrontations that take up time (Johnson & Johnson, 2014). This impedes the collaborative decision making process and voids one of its major features: unanimous backing of the decision. A pilot study can be used to estimate a timeframe for the discussions that is long enough to not impede the collaborative decision making process but short enough to be feasible for the study.

11. Future research

An indication of an effect of the use of the app for group related items was found, but not for the individual items. This observation may be investigated in a study specifically designed to test this in order to determine if an effect is present. The possible improvements mentioned should be taken into account to ensure the results are valid. This support system was not designed for either level of support specifically, but appeared more effective for group related processes than for individual processes. It would be interesting to find out if other general support systems for meetings or discussions show similar patterns of effectiveness. Consequently, the factors leading to general versus individual effectiveness can be identified. This knowledge can then be used to specifically target effectiveness during the development of support systems.

The use of the app showed a trend towards an effect on the focus on the topic at hand during the discussions but this was not confirmed by objective measures in this study. It was hypothesised that the awareness of relevant and irrelevant topics did indeed increase, but after they occurred. The participants can then swiftly move on if a certain topic is recognised as irrelevant. This would lead to a subjective increase in focus, but not in an objective increase as was defined in this study. A study researching the differences between the subjective focus that was indicated by the participants and the objective focus can make clear what effect the app had exactly. The hypothesised explanation for the discrepancy between the subjective and objective measure can be tested if different objective measures are used, or, if a more qualitative approach is taken, by asking the participants in what way the focus improved.

The discussion topic had a significant effect on the clarity of communication flow and an indication of an effect of the topic depth, discussion effectiveness and on the group related factor were found. It was already known that discussion topic can influence the effectiveness of a discussion support system and the discussion itself. However, an indication of an effect was found on group related items only, not on the individual aspects of the discussion. It would be interesting to find out if this trend is observable for other support systems.

This study has tested the effectiveness of the app on groups without maturity, which influences the ability of the group to perform with full effectiveness. The effects that the use of the app have might be different with groups that have previous experience with working together and having discussions. A group might be able to make better use of the app which would improve its effectivity. On the other hand, the app might only have impact because the group had not reached maturity yet, which might make the app less useful in mature groups. Information about the effectiveness of the

app in groups of differing maturity is of critical importance for the application of the app and for possible target groups and is therefore an interesting topic for future investigation.

12. Conclusions

The results from this study show an indication of an improvement of focus on the topic at hand when the app is used as was hypothesised in hypothesis 2. None of the factors that were thought to be influenced by the use of the app were directly and significantly influenced. However, the focus on the topic, the clarity of communication flow, the discussion effectiveness, and the amount of structure had a significant predicting power for factor 1, which might be influenced by the app. The corresponding hypotheses 2, 5, 6 and 11 cannot be confirmed as they were hypothesised but the effect of the use of the app on focus seems likely and the effect of the use of the app on clarity of communication flow, discussion effectiveness and the amount of structure is plausible if the app does indeed influence the group related items. The remaining hypotheses 1, 3, 4, 7, 8, 9 and 10 cannot be confirmed in any way, as there is no indication for a direct or indirect effect as a result from this study. An indication of an effect of the discussion topic on the discussion effectiveness and the depth of the discussion were found, as well as a significant effect of the topic on the clarity of communication flow. The effect that the use of the app and the topic had was not significant for either factor 1 or factor 2 so no definite conclusion can be drawn regarding their effects on the factors. However, there is an indication that an effect of the use of the app and the topic on factor 1 might be present, which means the app might be useful in improving group processes.

The results from this study indicate that the focus on the topic at hand might improve by using the app and it was hypothesised that this is due to the awareness of the comment types. This awareness may be a foundation for a support system and could provide a useful addition to the already available discussion support systems.

13. References

Baumeister, J., & Striffler, A. (2015). Knowledge-driven systems for episodic decision support. *Knowledge-Based Systems, 88*, 45-56. doi:10.1016/j.knosys.2015.08.008

De Winter, J. C. F., & Dodou, D. (2010). Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research and Evaluation, 15*(11).

Johnson, D. W., & Johnson, F. (2014). *Joining Together*. Edinburgh Gate, Essex: Pearson Education Limited.

Kolfschoten, G. L., Lukosc, S., & Leimeister, J. M. (2013). Introduction to the Intelligent Collaboration Support Systems Minitrack. 490-490. doi:10.1109/hicss.2013.331

Lehto, M. R., & Nah, F. (2006). Decision-making Models and Decision Support. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics, Third Edition.*: John Wiley & Sons, Inc.

Miller, S. (2016). Health condition monitoring system. Retrieved from https://www.utwente.nl/en/organization/news-agenda/special/2016/living-smart-campus/2nd-call-projects/asphalt-road-health/

Owen, D. (2015). Collaborative Decision Making. *Decision Analysis, 12*(1), 29-45. doi:10.1287/deca.2014.0307

Poole, M. S., Holmes, M., Watson, R., & DeSanctis, G. (1993). Group Decision Support Systems and Group Communication. *Communication Research, 20*(2), 176-213. doi:10.1177/009365093020002002

Razmak, J., & Aouni, B. (2015). Decision Support System and Multi-Criteria Decision Aid: A State of the Art and Perspectives. *Journal of Multi-Criteria Decision Analysis, 22*(1-2), 101-117. doi:10.1002/mcda.1530

Reinig, B. A., & Mejias, R. J. (2014). On the Measurement of Participation Equality. *International Journal of e-Collaboration, 10*(4), 32-48. doi:10.4018/ijec.2014100103

Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision Support Systems, 3*, 111-126.

Todd, P., & Benbasat, I. (1991). An Experimental Investigation of the Impact of Computer Based Decision Aids on Decision Making Strategies. *Information Systems Research, 2*(2), 87-115.

Todd, P., & Benbasat, I. (1992). The Use of Information in Decision Making: An Experimental Investigation of the Impact of Computer Based DSS on Processing Effort. *MIS Quarterly, 16*(3), 373–393.

Weiser, M. (1991). The Computer for the 21st Century. *Scientific American*.

Xinghang, J. D., Nada, M., & Guillaume, D. (2015). *cooperative knowledge representation and classification: application for design projects*. Paper presented at the IJCAI, Buenos Aires.

Yates, J. F., Veinott, E., & Patalano, A. L. (2003). Hard decisions, bad decisions: On decision quality and decision aiding. In S. Schneider & S. J. (Eds.), *Emerging Perspectives in Judgment and Decision Research*. New York: Cambridge University Press.

14. Appendixes

Appendix A

The questionnaire administered from each participant after each discussion.

Hello, **Discussion 1/2**

Thank you for filling in this questionnaire. It consists of eleven questions where you should indicate the statement that fits most with your feelings regarding the question. Indication should be made by filling in the circle above the statement.

Sona number: _____ **Used app this discussion:** O Yes / O No

How many contributions (arguments, counterarguments and proposals) did you feel like you have made?				
O *Very little*	O *Little*	O *Not very little; not very much*	O *Much*	O *Very much*
What was the quality of the decision made at the end of the discussion?				
O *Very poor*	O *Poor*	O *Not poor; not well*	O *Quite well*	O *Very well*
What was the quality of the discussion in general?				
O *Very poor*	O *Poor*	O *Not poor; not good*	O *Quite good*	O *Very good*
How effective was the discussion?				
O *Not at all effective*	O *Not very effective*	O *Not ineffective; not very effective*	O *Quite effective*	O *Very effective*
How clear was the communication flow?				
O *Not at all clear*	O *Not very clear*	O *Not unclear; not clear*	O *Quite clear*	O *Very clear*
How structured was the discussion?				
O *Not at all structured*	O *Not very structured*	O *Not badly structured; not well structured*	O *Quite well structured*	O *Very well structured*
How focussed was the discussion? (Was it on-topic?)				
O *Not at all focussed*	O *Not well focussed*	O *Not badly focussed; not well focussed*	O *Quite well focussed*	O *Very well focussed*
How clear was the content of the discussion for you?				
O *Not at all clear*	O *Not very clear*	O *Not unclear; not clear*	O *Quite clear*	O *Very clear*
What was the depth of the discussion?				
O *Very shallow*	O *Shallow*	O *Not shallow; not deep*	O *Quite deep*	O *Very deep*
How much did you participate in the discussion?				
O *Very little*	O *Little*	O *Not little; not much*	O *Quite a lot*	O *Very much*
How much collaboration took place during the discussion?				
O *Very little*	O *Little*	O *Not little; not much*	O *Quite a lot*	O *Very much*

Appendix B

Descriptive statistics for the subjective measures.

Descriptives			
		Statistic	Std. Error
Amoutn of contributions	Mean	3,13	,152
	Variance	,716	
	Std. Deviation	,846	
	Skewness	-,258	,421
	Kurtosis	,336	,821
Decision quality	Mean	3,35	,127
	Variance	,503	
	Std. Deviation	,709	
	Skewness	-,644	,421
	Kurtosis	-,709	,821
Discussion quality	Mean	3,26	,113
	Variance	,398	
	Std. Deviation	,631	
	Skewness	-,252	,421
	Kurtosis	-,515	,821
Discussion effectiveness	Mean	3,42	,111
	Variance	,385	
	Std. Deviation	,620	
	Skewness	-,569	,421
	Kurtosis	-,507	,821
Clarity of communication flow	Mean	3,58	,159
	Variance	,785	
	Std. Deviation	,886	
	Skewness	-,568	,421
	Kurtosis	-,378	,821
Amount of structure	Mean	2,94	,153
	Variance	,729	
	Std. Deviation	,854	
	Skewness	,472	,421
	Kurtosis	-,599	,821
Focus	Mean	3,90	,134
	Median	4,00	
	Variance	,557	
	Std. Deviation	,746	
	Skewness	-,868	,421
	Kurtosis	1,419	,821
Clarity of content	Mean	4,32	,097
	Variance	,292	
	Std. Deviation	,541	
	Skewness	,105	,421
	Kurtosis	-,668	,821
Depth	Mean	2,94	,103
	Variance	,329	
	Std. Deviation	,574	
	Skewness	-,015	,421
	Kurtosis	,336	,821
Participation	Mean	3,19	,150
	Variance	,695	
	Std. Deviation	,833	
	Skewness	-,389	,421
	Kurtosis	,571	,821
Collaboration	Mean	3,58	,111
	Variance	,385	
	Std. Deviation	,620	
	Skewness	-,326	,421
	Kurtosis	,025	,821