SOCIAL MIRRORS: VISUALIZATION AS CONVERSATION FEEDBACK

BY

ANTHONY DAVID BERGSTROM

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Associate Professor Karrie Karahalios, Chair, Director of Research
Associate Professor Brian Bailey
Professor Marshall Scott Poole
Professor Gary Olson, University California, Irvine
Abstract

Social cues shape our interactions and the way others see us. We convey a persona through our actions and words, but we are not always aware of how we are seen and interpreted. We use cues to coordinate conversation, determine speaker, and sense the mood of a group. However, online social spaces have expanded the ways in which we interact with others. Due to the nature of online communication, interactions persist in a manner that allows us to review our action. Otherwise devoid of traditional social cues, online places offer different mechanisms for mediation and feedback that are often not available in face-to-face conversation.

This dissertation applies affordances of computer mediated communication such as persistency, anonymity, large group backchannels, and archival into realized conceptions of a Social Mirror — a visualization to augment real-time conversation with additional feedback. The Social Mirrors capture interaction patterns and reflect visually aesthetic representations in real-time. The designs demonstrate additional cues alongside the traditional social cues one might expect by integrating this feedback directly into the conversation space. I show how this feedback influences the perception of self to change interaction, encourage self-evaluation and enable participants in conversation.

In the first chapters of this work, I establish that feedback from the social mirror affects people differently based on what the social expectations dictate. These mirrors, and those in related work, consistently demonstrate a balancing of conversation. This is not because the visualizations direct an individual to change; rather, the participants shape themselves into a perceived balance with each other through the mirror. In a mirrored conversation, the talkative back off and the quiet speak up while reporting their contribution was respectively dominating the social mirror and lacking from the social mirror.

More important than the individual projects, the concept of the social mirror provides a tool for group interaction. With social mirrors, groups gain access to new cues in real-time, but can continue to mediate interaction through self-directed means. The underlying theoretical construct I contribute
(presented in Chapter 4) defines the characteristics of social mirrors leveraged in this work: capturing interaction data, visualizing this data, and encouraging reflection on this data. The four example social mirrors — Conversation Clock, Conversation Votes, Fragmented Social Mirror, and Conversation Clusters — illustrate the social mirror as applied to conversation, however social mirrors can have broader applications that I hope contribute with future developments in other domains.

Specifically, this dissertation presents five contributions. First, I put forth the Social Mirror as a theoretical construct. I describe the necessary characteristics based in social visualization with influence from signaling theory, group dynamics, and accommodation theory. Social mirrors capture interaction, visually construct a cue, and feed that information right back into the current interaction and are not limited to my own designs. It is a feedback loop of self-evaluation that creates otherwise inaccessible social cues. Second, I present my visualizations for their design contributions. Each mirror combines abstracted audio visualization and augmented spaces with informative and aesthetic imagery. Each visualization leverages design to inform the viewer in an environment that is already visually demanding attention. In order to convey meaning, I rely on a well designed structure with minimal distractions. In this dissertation, I present the unique design challenges and visual solutions to ensure that the abstracted audio is a beneficial augmentation to conversation. Third, I demonstrate that the presence of a visualization encourages a “balanced” conversation. This is not because of the demands of the visualizations, but because the participants shape themselves into a perceived balance with each other through the mirror. I demonstrate this link initially with the Conversation Clock in terms of contribution: the talkative back off and the quiet speak up. This effect persists with the addition of approval/disapproval feedback in Conversation Votes. Most importantly, a greatly distorted visualization has relatively little effect on this balance. The presence of these social mirrors and reminder of a person’s activity is more than enough to encourage “balance.” Fourth, I demonstrate that explicit but anonymous feedback can empower individuals in conversation. Rather than refraining from participating, participants use the anonymous backchannel to express their opinion without the need to speak up. This channel leads to more speaking and more student initiated questions in a classroom setting. Finally, I present a social mirror that demonstrates how the context of conversation can be detected and incorporated into the social mirror. Accounting for poor speech recognition, the Conversation Clusters mirror allows a social mirror timeline to be automatically annotated by topic for longer archival and review. As a whole, this dissertation presents the social mirror as an interface to improve upon self reflection in conversation. Each project presented illustrates the use and benefit of such mirrors and opens the doors to additional mirrors by others.
to Lin
Acknowledgments

Thank you, Karrie, for the exciting times in Social Computing. I have always enjoyed your eclectic interests in new technology and openness to exploring new ideas. Your guidance has lead me to incorporate many areas outside of Computer Science in a way that drives me to be more open to new research possibilities and ideas that may otherwise be overlooked.

Special thanks are extended to my committee — Brian Bailey, M. Scott Poole, and Gary Olson — for their guidance as committee members and their advice in my research. I am thankful for their perspectives on my work.

I could not have finished without the support of many friends and family. Though all members of the Social Spaces group were tied to my experience I must recognize Ankur and Mat as the firsts. Mat, I would additionally recognize for forcing me to act on my creative ideas rather than just letting me talk about them. My long-term office-mate, Eric, always listened to my craziest of ideas and helped me through my most frenzied states. I could always count on his confidence in my abilities when my own faltered. Jake and Shamsi provided excellent support as the senior HCI graduate students, I only hope I did as much for those following me. My apartment-mate, Jacob, deserves recognition for our many years of nonsensical conversations, celebration of midwestern values, and shared suffering. I also recognize Illini Jujitsu and Jeff as being the most central to my physical well-being. Additionally, to the many unnamed friends and family, thank you.

Though I may have spent more time at Siebel Center than in my apartment, my stipend ensured a level of minimal comfort. First, I thank the taxpayers of Illinois, who paid me to assist in the teaching of their undergraduate children. I thank Samsung and the National Science Foundation for funding me through Karrie’s research assistantships. A special thanks goes to Thomas Siebel and Siebel Scholars Foundation for recognizing me and for helping me through at the end.

Finally, I thank my wife, Lin. Of everything I have gained along the way, you surpass the rest by far. I look forward to our years ahead.
# Table of Contents

## List of Tables

<table>
<thead>
<tr>
<th>Table of Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
</tbody>
</table>

## Chapter 1 Introduction

1.1 Scope ........................................... 1  
1.2 Contributions ................................... 2  
1.3 Overview of the Dissertation ................. 3

## Chapter 2 Theory

2.1 Visual Information Design ..................... 6  
2.2 The Social Self ................................ 10  
2.3 Groups and Their Dynamics .................... 13  
2.4 Visual Conversation ............................ 18  
2.5 Ambient Displays and Augmented Spaces ...... 21

## Chapter 3 Social Mirrors

3.1 Foundations ..................................... 25  
3.2 Hall of Mirrors ................................ 26  
3.3  
3.4  
3.5  | 27 |
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>These works demonstrate collocated conversation visualizations focused on changing group dynamics.</td>
<td>21</td>
</tr>
<tr>
<td>4.1</td>
<td>Results of the Likert Scale surveys in the Conversation Clock user study. Using a Wilcoxon Signed Rank Test, we show that participants reported a greater effect of the long history than the short history.</td>
<td>41</td>
</tr>
<tr>
<td>4.2</td>
<td>For analysis, we hand-coded video taken from each session for gaze direction and gestures over the table. Using a repeated measures ANOVA, we found significant changes in the gaze direction of participants. A notable, but not significant, change in table gestures was also observed.</td>
<td>41</td>
</tr>
<tr>
<td>4.3</td>
<td>Audio data, automatically logged and analyzed, was available for each participant. Expecting different changes to occur in each, we categorized participants as talkative or quiet based upon the initial session. We ran a repeated measures ANOVA and found talkative participants change the length of their turns while quiet participants changed the number of turns.</td>
<td>42</td>
</tr>
<tr>
<td>4.4</td>
<td>Re-sorting to indicate the first condition seen and the second condition seen, it becomes apparent there were novelty effects. Participants in both the talkative and quiet participation categories altered the amount of time spent leading. Examining the other notable statistics, the talkative category seems to have altered the length of their turns while the quiet participants seemed to change the number of turns to accomplish this feat.</td>
<td>43</td>
</tr>
<tr>
<td>5.1</td>
<td>The above table demonstrates the impact of the visualization sessions across all participants. In general, there was very little change to be attributed to the condition alone. The exception being a decrease in turn length during the final session.</td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td>The above table summarizes the results of the Likert scale survey in the second session. We see that participants reported being somewhat altered in their interaction and noticing alteration in others. The majority of these questions were phrased as 1-Disagree to 7-Agree.</td>
<td>56</td>
</tr>
<tr>
<td>5.3</td>
<td>Splitting the participant groups on three different variables — Talkativeness, Voting Patterns, and Voting Support — the linear mixed effects model highlights the differing interaction emerging from our split categorizations over time and across the visualization conditions. Extended versions of this table appear in Appendix B.</td>
<td>57</td>
</tr>
</tbody>
</table>
54 individuals rating our icons produced these icons as the most indicative of their intended messages. The results informed our final selection of icons seen in the interface (Figure 6.3). One other message, to indicate the speaker was moving too fast, produced no positively rated icons. All bar graphs appear at the same scale.

This table shows the number of people in the classroom for each session and the number of visible computers. Though we requested that students bring their computers to the sessions with the Fragmented Social Mirror, the number of computers remained essentially unchanged.

Students reported they were uncomfortable asking questions in class, but more comfortable asking questions in smaller recitation sections. They are aware that they do not participate in class and they do not ask the questions that they have. There is a reluctance amongst the students to speak up and participate. 23 students responded to our survey; all graphs have the same scale.

The Fragmented Social Mirror encouraged students to initiate questions and dialog with the instructor. Many dialogs began with or contained questions while others were responses to indicate comprehension of the material. The anonymity of the interface also encourages unrelated dialog in class. A dialog in the Fragmented Social Mirror could include multiple messages — thus, # dialogs (# messages) includes both the count of interaction instances and total messages.

Students reported the Fragmented Social Mirror encouraged participation and connected them to the lecture. Students reported the lectures being more enjoyable and saw the feedback as a worthwhile addition to the lectures. All graphs appear to the same scale, 16 students responded.

A repeated measures analysis comparing Talkative/Quiet participants, emphasis strategies, and Talkative/Quiet participants across the emphasis strategies. Extended versions of this table appear in Appendix E.

The above table shows the results of a repeated measures analysis comparing between conditions, the Talkative/Quiet split, and the two combined. Extended versions of this table appear in Appendix E.

This table shows differences of the visualization's effects on Talkative and Quiet participants.

This table shows differences of the visualization's effects on Active Voters and Less Active Voters.

This table shows differences of the visualization's effects on Heavily Supported and Lightly Supported participants.

Collocated groups, contrasting Talkative and Quiet participants

Collocated groups, contrasting states of emphasis

Collocated groups, contrasting Talkative and Quiet participants in states of emphasis

Remote pairs, contrasting Talkative and Quiet participants

Remote pairs, contrasting states of emphasis

Remote pairs, contrasting Talkative and Quiet participants in states of emphasis
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Chat circles in [Donath et al., 1999] create an artificial cocktail party. To converse, individuals must be near each other. Additionally, the full history is not accessible in order to encourage live participation.</td>
</tr>
<tr>
<td>2.2</td>
<td>The Babble interface in [Erickson and Laff, 2001] shows who is active in a chat room. Individuals towards the center are more likely to read and respond to messages while inactive people appear outside the circle.</td>
</tr>
<tr>
<td>2.3</td>
<td>From a personal repository of emails and text, Themail in [Viégas et al., 2006] is able to present a compelling timeline that hints at stories in the life of the user.</td>
</tr>
<tr>
<td>2.4</td>
<td>Visiphone in [Donath et al., 2000] connects two remote locations with a constant connection. The audio signals produce a visual stream of circles one side in blue one side in orange.</td>
</tr>
<tr>
<td>2.5</td>
<td>Work in [DiMicco et al., 2004] demonstrates the use of a wall projection and histogram to demonstrate participant contribution. Labels of &quot;under-participating,&quot; &quot;participating,&quot; or &quot;over-participating&quot; characterize each individual.</td>
</tr>
<tr>
<td>2.6</td>
<td>Research in conversation visualization includes the exploration of tabletop displays, wall displays, and mobile displays. Though the visualizations differ, commonly there is a tendency to balance conversation amongst all participants.</td>
</tr>
<tr>
<td>2.7</td>
<td>[Leshed et al., 2009] fosters a more positive environment by highlighting agreement and discouraging negative words in discussion.</td>
</tr>
<tr>
<td>3.1</td>
<td>The image of the Conversation Clock presents a circular representation of conversation. Each bar a moment of time, the captured aural history defines the concentric rings. The visualization highlights characteristics such as domination, silence, and turn-taking.</td>
</tr>
<tr>
<td>3.2</td>
<td>The initial study with the Conversation Clock began explorations into the use of feedback, emphasis, and content incorporated into the social mirror.</td>
</tr>
<tr>
<td>4.1</td>
<td>Here we see some of the initial design iterations. Each design underwent informal testing to determine the most useful aspects and future designs. Groups would sit at the table for brief conversations. The design was iterated weekly and refined into the Conversation Clock.</td>
</tr>
<tr>
<td>4.2</td>
<td>The earliest Conversation Clock designs featured microphones embedded directly into the table itself. This design offered a more natural &quot;just sit down and chat&quot; design. However, lapel microphones provided a better signal.</td>
</tr>
<tr>
<td>4.3</td>
<td>The Conversation Clock demonstrates simple patterns of conversation based on microphone input.</td>
</tr>
</tbody>
</table>
4.4 Above are examples of overlapping speech, silence, turns, and other characteristics shown in the Conversation Clock. 

4.5 In the above images, we see how the Conversation Clock structures samples over time, clockwise around the circle. When a ring is filled, the visualization compresses to make room for the next minute. 

4.6 The Conversation Clock increased the time spent looking away from others in conversation. Most of the looks to the Conversation Clock were in addition to the baseline set in the non-visualization sessions. 

4.7 The Conversation Clock tended to reduced the number of gestures made by participants during conversation. 

5.1 The Conversation Votes table allows participants to view their conversation with an additional perspective. All members at the table are able to indicate their agreement or disagreement. 

5.2 The central timeline dominates the Conversation Votes visualization. Each second, it samples and appends the current state to the leftmost end of the progression. All bars in the progression slide to the right to make room. Dotted rectangles mark the time a vote took place and multicolored bars show any simultaneous speech. 

5.3 On either side of the central timeline a longer history builds up. Each collection of samples represents one minute of time, the most recent picture furthest to the left. Over the course of a minute, the left most collection starts with no samples and grows. It then shifts to the right when completed. 

5.4 The voting button is a handheld plate with a colored circle indicating where to push. The button allowed for discreet and hidden voting. The button provided no tactile feedback, a light press was sufficient to vote and alter the visualization. 

5.5 The effect of a positive vote lengthens the otherwise uniform bars. 

5.6 Though the two groups behaved similarly without the visualization present, Figure 5.6a shows active voters increased their contribution while less active voters decreased during the visualization condition. Figure 5.6b shows that the most active voters did not feel representation was comparable until the visualization was present in the second mini-session. 

6.1 Many of the early prototypes for a large group continued to display captured audio in the display. In a large unbalanced conversation such as a lecture, that feedback seemed less important to a participant. We opted to drop this feature in favor of a simpler interface. 

6.2 This prototype allowed participants to send feedback to any speaker and not just the lecturer. However, initial testing deemed this interface too involving. Participants would dedicate much more attention to the interface than to the interaction. However, if there were fewer cues, as in a teleconference, it may still prove beneficial. 

6.3 The input interface provides a simple design for a small window. It allows a student to send one of the four preselected messages and text (when applicable) from their computer to the public display. It does not require a lot of attention in order to encourage more attention on the classroom activity.
6.4 We projected this larger display at the front of the classroom while the lecturer could see
a smaller personal copy. The screen anonymously captures questions and comments
that students might not otherwise ask. The public display groups messages by icon
and highlights the most recent feedback from the class. To the left of each icon, a
counter indicates multiple messages of the same type, particularly useful when using
the yes/no feedback buttons. As we saw in our pilot study, it enables the students to
initiate a dialog with the lecturer.

6.5 We tested the Fragmented Social Mirror in a large lecture hall with three projection
screens. Typically, the instructor repeated the same material on all three screens. For
the study, the Fragmented Social Mirror replaced one screen during class-time. The
lecturer also has a private screen showing the public display.

7.1 For the same conversation, the above images depict renderings for the blue speaker
under (A) Normal, (B) Speed Emphasis, (C) Amplitude Emphasis, (D) Brightness

7.2 The results of the pilot study indicate all strategies encouraged overestimation with at
least one setting (Low, Medium, High). The highest bar in each represents a significant
deviation from the control.

7.3 Talkative and Quiet people react differently to the presence of the visualization. How-
ever, throughout the conditions only AM differed from the control condition for
both Talkative and Quiet people.

7.4 People perceived that they were talking less and more than usual in parallel to their
emphasis conditions. The Quiet participants were more in tune with the emphasis,
whereas the Talkative generally perceived themselves as more subdued in the presence
of the visualization.

7.5 The remote study demonstrated based on emphasis condition played a significant
role in contribution to conversation (F(2,48) = 4.2, p < 0.02). The change between
emphasized and underemphasized

7.6 The average contribution under each condition produce generally similar lines. The
exception is at minutes 6–8. See Figure 7.7 for more details.

7.7 A minute by minute look at the contribution of participants shows that most of the
difference in condition occurs at minutes 6–8. In fact, comparing these three measure-
ments show that they are significantly less than all other minutes (t(30) = 3.08 ,
p < 0.002)

8.1 An example Query to generate related concepts from Wikipedia using ESA, described
in [Gabrilovich and Markovitch, 2006a].

8.2 Captured conversation clusters on the table in real time. The clusters reveal the current
model of conversation as the computer sees it. Individual people can correct this
model directly.

8.3 The thread history visualization demonstrates the progression and change of conver-
sational topics.

8.4 To calculate the similarity and distance of two topics.

8.5 Recalculating topic and words weights when combining two topics.

8.6 The transcript is divided into segments that are then grouped into bins based on
chronology. Each bin produces a consolidated set of topics that are connected to
similar topics in other bins based on their similarity. These connections define the
themes that permeate conversation.
8.7 Choosing the words to represent a thread of conversation: the highest ranked word that is current and unused.

8.8 This conceptual image depicts the goal of the Conversation Clusters project. It shows a simplified depiction of conversation that is easy to pick out the topics of conversation. In addition to showing topics, it also depicts the relative importance of each thread by size and placement.
CHAPTER 1

Introduction

"We must develop tools that people prefer to use even when they have the option of interacting as they have heretofore in physical proximity. We must develop tools that go beyond being there."

Hollan and Stornetta, Beyond Being There (1992)

Social Visualization reveals the dynamics of social groups based on the traces they leave behind. Properly constructed, these visualizations allow a viewer to easily understand a social place. They may generally be geared toward the external explorer, but social visualization is in a prime position to inform the same groups it models during interaction. I show how social visualization empowers new voices in conversations and raises awareness of interaction patterns by using a construct I dub Social Mirrors. The social mirror leverages the natural human tendency to control how others perceive them to encourage a reflective self moderation. My work is based in the domain of conversations, but the social mirror can be applied more widely.

When we talk with others, there is an outward public persona that presents only a glimpse of one’s personality. This is by design — we care how others perceive our actions and change personas to fit the context of a situation. We watch others to understand how we are seen, but we have only a vague notion as to how we are seen by them. Does this change when we augment conversation? How do we react when our ephemeral actions become persistent in real-time? Can we use this persistent channel and awareness to enable other voices in conversation? The purpose of this dissertation is to answer these questions using social mirrors as personally enabling visualizations.

Conversation, meetings, and face to face social interaction are historically well studied in fields such as sociology, psychology, philosophy, economics, and more (Canfield, 1990, Cleveland, 1888). The interaction rituals and patterns that have developed into culture allow us to confidently navigate a broad variety of social conversations without too much trouble. We change our tone, our word choice, our facial expressions to meet these expectations. All the while, we pay attention to our audience for feedback and signals that indicate how we are being perceived. For example, a sensitive topic might be breeched delicately. We offer many opportunities to back out of the topic and pay close attention to the responses. If it appears we are beginning to press too hard, we might chose to back off and leave topic unexplored.
This collection of works seeks to augment the face-to-face communication channels by applying useful cues from online mediated environments directly into real-time conversations. As technology permeates everyday life, signals in communication adapt to fit the new medium based on what is available and possible. Some signals parallel face-to-face interactions, while others allow new conventions that work only in a mediated environment. For example, online groups in chatrooms and forums can include a far larger set of people with mechanisms to indicate approval, to preserve anonymity, to archive textual conversation, and to enable more people to directly interact across traditional barriers. The ease of communication and the flattening of access has led some to characterize the internet as a leveling ground where any voice can be heard with fewer repercussions to their identity or reputation. As a simple example, email allows employees to break hierarchal power barriers in the workplace because it offers an easily accessible channel that is less personal (Sproull and Kiesler, 1991).

Social mirrors provide a new channel in face-to-face augmented spaces that enable self-motivated change. A person viewing the social mirror can, and is, influenced to meet their own preconceived notion of how they should be seen by others. The mirror serves as a visual cue that spurs a person to act. Though social mirrors are effective at influencing conversation, the relation between depiction on the social mirror and degree of influence is not directly proportional. A highly manipulative visualization is only slightly more effective than a true visualization. This highlights the effect of self moderation and indicates the importance of internal motivation to the use of social mirrors in shaping participation. In conversation, I use the social mirrors to enable underrepresented participants in conversation. In addition to encouraging a more balanced conversation between the talkative and quiet, I show the use of explicit anonymous feedback via new back-channels to conversation shows that individuals can effectively use these channels to assert themselves into conversation. Face-to-face interactions are constrained to a limited number of people who can speak and influence the direction of conversation at any given time. My conception of the social mirror opens the conversation to more voices participating — be it a quiet person struggling to get a word in, a shy person afraid to speak up, or an individual lost in a crowd of others with no opportunity to contribute. Social mirrors provide an explicit awareness channel to spur individual action and change.

1.1 Scope

My work focuses on raising the awareness of a person's social appearance using visualization as feedback in small group conversations. The bulk of this dissertation examines groups of size 3–4 in collocated casual conversation. This simplifies the conversation as the group is unlikely to splinter into
side conversations or individuals move between cliques; however, the small size provides a microcosm of any casual group dynamic. As larger conversations tend to splinter into groups of size 3–5, many of the results in my dissertation could be applied to these splinter groups.

This dissertation does not cover every style of group and conversation under the influence of social mirrors. While size is probably one of the most influential characteristics, the power dynamics, the purpose of meeting, distributed locations, and cultural context also impact the way interaction unfolds. Though my work can inform these other areas by framing them as group evaluation of individuals, these areas must remain potential future explorations. With the exception of a single remote study and a brief application of feedback on a large classroom, I do not attempt to directly explore these other areas in my studies.

1.2 Contributions

This thesis makes the following contributions:

1. **The theoretical construction of a social mirror to describe and design new interfaces**: Chapter 3 outlines the tenets of social mirroring of visualization as feedback into the social environment it monitors. Each experiment and project in this dissertation fleshes out different aspects of social mirroring: persistent presentation, group perception, and capturing details from interaction. Though I focus on conversation as a domain, the ideas of social mirroring are extensible to other areas of interaction and ripe to be applied.

2. **The design process and product that lead to informative and aesthetic designs for active environments**: Each visualization is the product of many prior designs and design decisions throughout the development process. I demonstrate prior designs and articulate how the final designs were chosen from that process. These designs are intimately married to their environment and compete for attention with the normal cues of conversation. In spite of the constraints of environment, the visualizations proved to be both informative and aesthetic.

3. **An analysis showing that the presence of visualization encourages balanced conversation**: Visualization of contribution in collocated conversation encourages the talkative individuals to speak less and the quiet to speak more. I show this initially using the Conversation Clock in Chapter 4 as have other researchers in their own experiments (detailed in Section 2.4). However, I continue to show similar results even in the presence of additional silent channels and distorted visualizations. Chapter 5 highlights the use of a silent channel to identify and encourage individuals
making beneficial contributions to conversation, however the balancing effect still remained. Stronger support is seen in Chapter 7 in which the balancing effect of a visualization still occurs in the presence of a highly distorted visualization. The distortion has some, but limited, effect on conversation. Even with the distortion, the balancing effect is still observed.

4. **Findings that show the benefits of anonymous feedback in collocated spaces**: I built two interfaces to focus on anonymous feedback and capturing group perception with Conversation Votes and the Fragmented Social Mirror in Chapters 5 and 6. Each of these works show that anonymous feedback is effective at providing insight into the rest of the group. It shows if others agree and understand points made in discussion, and it encourages initiative in questioning. Perhaps more than the other experiments, the use of anonymous feedback was effective at empowering individuals to speak, participate, and express their views in an environment that they showed initial reluctance to disturb.

5. **Applications to capture and characterize conversation context**: Each project in this work captures aspects of conversation such as aural participation, group perception, new question, or as in Chapter 8 the thematic progression of conversation. In light of the context, people report wanting to change their participation — for example balancing conversation as above. The final contribution presented in this work demonstrates that a detailed depiction of thematic topic development can be captured in real time and tied to the social mirror for archival use and summary.

### 1.3 Overview of the Dissertation

This dissertation presents valuable channels based on real-time interaction. The underlying goal is to enable underrepresented opinions, shy people, or a large audience by making the social-dynamics of a group more apparent. The use of these interfaces demonstrate the capacity to augment a collocated interaction. The use of digital cues incorporated into real-time domains has much to offer in terms of useful feedback.

In Chapter 4 I discuss the background work that inspired and informed my own work. I situate my work using classic Goffman, Signaling Theory, and Accommodation Theory as motivators for individuals to change their participation. I discuss the progression of group interfaces that has influenced and augmented group interaction. I also present a number of varied work in social visualization, ambient displays, and augmented spaces that have developed in the Human Computer Interaction community to investigate related questions.
In Chapter 3, I present my definition and conception of a Social Mirror. This umbrella term collectively summarizes my own work and could include a number of past works by other authors. I argue that social mirroring is a subset of social visualization and group interfaces. This chapter draws from our prior publication discussing abstract audio visualization as a signal (Karahalios and Bergstrom, 2009). I also summarize each project presented in this thesis and how it fits the term social mirror.

The remaining chapters detail the experiments I conducted as a part of my research. Chapter 4 depicts the Conversation Clock and work published as a part of my master’s thesis (Bergstrom, 2006) and other publications (Karahalios and Bergstrom, 2006; Bergstrom and Karahalios, 2007a; Bergstrom and Karahalios, 2007b) that started my exploration of social mirroring and led to more questions.

Chapter 5 depicts the Conversation Votes interface as an attempt to explore the use of anonymous feedback in small group conversations. This work also appeared as a part of early design work on tabletop interfaces (Karahalios and Bergstrom, 2006; Bergstrom and Karahalios, 2007c) and again upon completion of the experiment (Bergstrom and Karahalios, 2009b).

Chapter 6 continues the exploration of anonymous feedback, but applies it to the classroom environment where the conversation is generally led by an instructor in the presence of a silent classroom. At the time of this writing, this work was accepted to appear in an upcoming conference (Bergstrom et al., 2011).

Chapter 7 returns to the Conversation Clock to further investigate the changes that we saw in chapters 4–6. At the time of this writing, this work is not yet published on its own, but appears as a technical report (Bergstrom and Karahalios, 2010).

Chapter 8 presents work to capture topics from conversation and model them as a part of the persistent social mirrors. Its design and implementation are also discussed in two short publications (Bergstrom and Karahalios, 2008; Bergstrom and Karahalios, 2009a).

Finally, Chapter 9 summarizes the broader impact and limitations.
My work demonstrates the ability of real-time visualization to heighten awareness and change personal contribution in conversation. In this review of literature, I first situate my work amongst the backdrop of visual design and Social Visualization — a method for visually understanding the traces that social interaction leaves behind. Social visualization demonstrates structure in data that reveals insights into the underlying social group. From my perspective, the work in social visualization generally and social mirrors specifically relies on the traditional ideas of face, the self, and signals between individuals. This literature establishes the basis for encouraging awareness of self when interacting in groups. The signals I construct with social mirrors stem from the natural signals that take place in conversation. I then discuss literature focusing on how group conversational dynamics are affected by different cues, signals, and group compositions. This leads into a subset of Social Visualizations that have focused specifically on the visualization of conversational cues, an area also situated in the same background of literature and most directly related to my own work. Finally, I end by discussing ambient and reflective interfaces intended to merge the computer interface into the less traditional displays to inform a user, and in the case of some interfaces, encourage deeper thoughts as to their own actions.

2.1 Visual Information Design

Visually structuring information allows us to externalize our knowledge and think differently about our situation. In much the same way as written language impacted culture and communication (Ong, 2002), the use of visualization in the form of graphics and charts impacts our ability to comprehend data (Tufte, 2001). Visualization allows for the simplification of large quantities of data to make the complex relationships apparent. It leverages our aptitude for visual input rather than raw data (Tufte, 1997).
The use of imagery to demonstrate data is a powerful tool, one that can be done well or poorly. The best design practices encourage showing data in a way that enables comparison. One might show data side-by-side or on in parallel, extraneous notation should be removed, labels should be present, and the information should take center stage (Bertin, 1983; Tufte, 2001; Tufte, 2006). With increased use of charts, graphs, and info-graphics daily life, people have become increasingly graphically literate. As a result, complicated visualizations and interactive exploration has moved beyond the traditional histograms and line graphs. However, many of the basics still apply.

The most important aspect of designing visual information is the viewer; it is the viewer's mind that makes the imagery meaningful (Arnheim, 2004). That imagery is only meaningful if presented in the proper context — a context that allows the viewer to perceive information and not just an image. When viewing an image, we are not merely perceiving it, we are giving meaning to the image. As a result, the whole of our design decisions impact the viewing and understanding of our data.

### Social Visualization

Social Visualization provides a visual reference for group dynamics often based in online interactions with limited communication channels. It presents a visual depiction of information that allows the viewer to recognize patterns in interaction data that show details about the underlying activities, social groups, and interaction.

Social Visualization can trace its roots to the development of social network diagrams. These diagrams made the connections between individuals explicit based on the observation of interaction (Freeman, 2000). As new social environments arose on computers, the lack of cues provided an opportunity to create new ones — not only for the purpose of recreating old cues, but to offer a framework for understanding the unfamiliar interactive space. The idea was to make online spaces and people differentiable from each other — to show an active social space versus an empty one. Erickson and Kellogg used the term "socially translucent" to describe an interactive space where people could be aware of others, accountable to others, and visible to others in the same space (Erickson and Kellogg, 2000). Visualization is one of the means to achieve such a goal.

Early work in social visualization explored the interactive patterns found in mailing-list, email, Usenet forums, and chat rooms. One project, Visual Who, offers a simple example showing how individuals group membership in mailings lists relates the users of a shared system (Donath, 1995). The visualization allowed mailing list anchors and a spring based system to spread the individual users around and characterize them. The results tended to cluster like-individuals together and allow for exploration.
Another mode created a more informed awareness display by allowing users to see only others who were currently logged in the system. However, to get at the more interactive domains like forums visualized in the projects PeopleGarden, Loom and later AuthorLines by showing the activity rather than membership of communities (Xiong and Donath, 1999; Donath et al., 1999; Viégas and Smith, 2004). All of these projects demonstrated interesting characteristics such as a broad user-base, dominant users, and dying communities. Of the three mentioned, AuthorLines specifically focused on the authors, rather than the characterization of the forum as a whole, by visually summarizing an authors posting patterns in a way that can be easily compared (Viégas and Smith, 2004).

![Chat circles in Donath et al., 1999](image)

In addition to descriptive visualizations, social visualization provides structure to interactions. Shown in Figure 2.1, Chat Circles simulates aspects of physical conversations into the online chat room. Participants had a distinct location on the screen which they could move through the chat-space. The location and proximity determined what activity was visible to the user. Neighboring comments appeared on the screen at the users location in a re-sized circle while distant activity appears without text. Thus, an individual can only converse with others who are nearby in the space (Donath et al., 1999). Another work showing activity in a chat room, the Babble interface keeps the standard chat system but augment that chat with an indicator of activity (Erickson and Laff, 2001). The most active people appear in the center of the circular visualization while less active appear at the edges or outside the circle. At a glance a user could infer whether and who might respond to a new message.

Figure 2.1: Chat circles in Donath et al., 1999 creates an artificial cocktail party. To converse, individuals must be near each other. Additionally, the full history is not accessible in order to encourage live participation.
Looking towards email, visualization reveals much more personal information. It offers a chance to learn not only about an online community of people, but to examine one's own personal data. Email being widely visualized domain, social network visualizations vary in both style and purpose. Building from social network theory, email captures a personalized social network diagram (Boyd and Potter, 2003) and the information transfer in a large company (Heer, 2010). Used in conjunction, these visualization strategies aid in the identification of cliques of individuals and diffusion patterns when investigating a company or group (Stasko et al., 2008). Aside from the full picture, email visualizations also focus on the reply structure of messages in order to navigate an inbox or archive (Kerr, 2003; Venolia and Neustaedter, 2003), while others focus on the content of the message by displaying the emotion inherit in the text (Viégas et al., 2004a) or picking out significant words (Viégas et al., 2006).

Social visualization has also taken to visualizing activity as a means to and encourage production and interaction. Visualizing group production and editing patterns in Wikipedia serves to highlight contentious articles or sections in articles (Viégas et al., 2004b; Suh et al., 2008). Similarly, demonstrating patterns in email and code contribution provides a method to understand open source development as a cohesive community (Gilbert and Karahalios, 2009).

Social visualizations capture and characterize the community they examine. It provides a snapshot of insight, they show the group dynamics of that community. In contrast, the social mirror described in this work continually updates to provide insight in the moment, and it's viewers make up the community being examined. Integrating into the conversation space, they create new channels to
Figure 2.3: From a personal repository of emails and text, Themail in (Viégas et al., 2006) is able to present a compelling timeline that hints at stories in the life of the user.

augment traditional social cues. They are presented such that viewers can evaluate their appearance and immediately respond to that appearance as they would any other form of social feedback. The cues convey additional information about the individual’s underlying true self.

2.2 The Social Self

The social mirror acts as a transitional object for the viewers. As defined by Winnicott, these objects exist between the psychic (internal) world and the external world (Winnicott et al., 1989). In some sense it is a part of the person. The cues visible on the table indicates some aspect of self, just as smiling or gesturing does. They extend a person’s identity to an object. However, it is also a part of the external world, separate from the individual.

This section discusses self by first reviewing the indefinite conceptions of “self.” The self varies with audiences and activity. The individual presents himself to others through the use of faces. Roles and expectations accompany face and offer a framework for interaction. Finally, communication itself occurs through signaling between two individuals, both acting through their chosen face. The self may have different faces for different individuals, but all of this is a part of the complexity of communication.

Self

The idea of “self” is a construction that people develop over time based on their interaction with others. It establishes the boundaries of where one perceives their identity ends and others begins. It
also establishes what parts of the physical world, for example our bodies and appearance, reflect on our own identity. As intuitive as it may seem initially, defining the “self” is not a simple task. Some might limit the self to the physical body, one’s own mental state, a continuous consciousness, or some other aspect of a person depending on the chosen perspective. Individuals develop their intuitive sense of self in childhood, as children begin to separate ourselves from others by observing [de Waal 1996]. By watching and interacting with other people children gain a sense that they are not the other person. Along the way, children learn the social protocols and cultural expectations of interaction. De Waal's work continues to argue that this self awareness cultivates empathy such that we care about how others feel. Allowing individuals to intuitively strive for mutually beneficial goals. Though from a philosophical standpoint the self may be an intuitive illusion, the model of a looking glass self – one in which we perceive ourselves as we think others do – influences human behavior [Canfield, 1990]. Whatever the philosophical position as to the internal illusion of self, studies show the setting, context, and roles impose on self identity changes the way individuals act [Yeung and Martin, 2003] [Ferraris and Carveth, 2003] [Haney et al., 1973].

In conversation, people make unconscious and subtle changes in different groups of people. Accommodation theory suggests individuals converge on a common interaction style during an exchange to show relative closeness [Giles et al., 1991]. Friendly circumstances encourage people to converge on similar grammar, accent, tone, word choice, etc to establish a cohesive group. As an example, a parent adopts a smaller vocabulary and grammar when speaking with a young child. The resulting simple demeanor shows closeness to the child. Conversely, differences in word choice, formality, and accent emphasize social distance. Speaking to one’s boss, individuals may adopt a professional vocabulary and mannerisms such as saying ‘sir’ or being overly attentive as a sign of respect, but this imbalance highlights the employee is not a part of the boss’s in-group [Giles et al., 1991] [Chaika, 1994]. The difference in politeness emphasizes the hierarchical nature of the relationship. However, studies show that age, role, and culture influence the degree to which a person accommodates in addition to closeness [McCann and Giles, 2006]. In each of these cases it is not that the individual is changing, but that they chose to present a different aspect of their self to others.

**Face**

Goffman refers to these different views of self as “face” [Goffman, 1963]. Face is something to protect and should show the best characteristics on an individual while hiding any negative characteristics. Social protocols and interactions that a child learns early on provide a guidebook on how to preserve
face in that culture (Goffman, 1967). People work to maintain a positive persona in the minds of those around them. As an individual, I work to ensure that friends, coworkers, and strangers perceive me as sociable, friendly, and competent. The face to which Goffman refers is more than just the expressions we have on the front of our head. In this context, physical and abstract social signals comprise face and influence others’ perceptions. In addition to those facial expressions, face is influenced by our choice of words, our actions, our charisma, and our deference.

As we interact, we have to be aware of how others perceive our actions in order to avoid embarrassment. In any role we take on, others hold expectations that we must meet. We change our behavior to fit the audience in order to make good impressions (Goffman, 1959, 1963). In the workplace, coworkers adopt casual styles with each other and make jokes, however those same employees can change drastically when others enter the office. The casual atmosphere can only last while it doesn’t reflect poorly upon the employees. If an important or perspective client were to make an appearance, social protocol becomes more formal and distant. The employees want to appear professional and hard working. The new atmosphere shifts to what the employees believe the customer wants to see. If a positive face is not maintained, discomfort enters into the social environment in the form of embarrassment or frustration (Goffman, 1967). The mutual desire to avoid such discomfort, encourages individuals to direct conversation away from events that might cause a loss of face.

**Signaling**

Signaling provides the basis for communication. Though human speech and writing are among the most information packed signals we use to relate information (Ong, 2002), communication signals are far more nuanced and use many more channels. Some information is better conveyed with non-verbal signals, as in the case of conveying human emotions and attitudes: researchers attribute only 7% of the signal to the words, 38% of the signal to voice and tone, and 55% of the signal body language (Mehrabian, 1971). As is the case when people hide their emotions, these signals can be in disagreement (Ekman, 2001). Thus, a person will not trust an easily faked channel as much as an unconscious one.

Signaling theory, from evolutionary biology, acknowledges both intended and unintended signals in which we hope to signal some feature. When speaking to the a friend, a person might intend to appear interested and engaged. Simply saying ‘yeah’ and ‘uh-huh’ shows this interest through back-channel information. Listeners use this back-channel without much thought, though the natural usage varies between cultures (White, 1989). However, unintended actions can betray a facade (Hauser et al., 2002). For example, if we happened to check the time during that same exchange, it might demonstrate the
desire to leave or maybe the tone of voice betrayed a desire to leave. It's left for the signal's receiver to judged the reliability of any signal they receive. Consider a well known example of rival channels in a restaurant. Common wisdom says that a person can more easily be judge by how they treat a waiter or maid contrasted with how they treat their dining companion [Jones, 2006]. The reason is that they are focusing on conveying the proper signals to their primary audience and often pay less attention to their appearance to others.

More explicit and directed signals can serve purposes other than social presentation. Signals can coordinate cooperative action. Organized groups, such as a baseball team or military unit, actively establish signals instead of speech. In this case, secretive signals allow for communication while preventing competitors from knowing the planned strategy. However, just working in the same environment on a regular basis is often enough to establish small signals that allow others to recognize cues to anticipate the needs [Heath and Luff, 1992]. Simply making our own work publicly available to our neighbor can be enough to encourage them to join and help. In complex work environments, the degree of actions taken via gesture is necessary to efficiently coordinate groups that require complicated orchestration.

To present the best face, a person must be aware of their own signals in all domains. As new forms of communication arise, new signals emerge. Research on social network sites demonstrates the active control of digital cues to present a desired image. In the dating world, these cues might be purposefully misleading in representing height, weight, age, and interests to attract a certain type of person [Hancock et al., 2004]. More general audiences in social networks provide cues through posted pictures, favorite music, movies, books, and recent activities to friends and acquaintances [DiMicco and Millen, 2007; Lampe et al., 2007]. The degree to which a person manages their identity varies by personality and goal. For example, whereas the average college student might be more open to sharing pictures and personal comments, a person seeking a job might choose to remain more professional in public dialog.

2.3 Groups and Their Dynamics

As with each individual, groups adopt set patterns of interactions based on context and audience. What may seem minor differences on the surface – group size, gender balance, location – can drastically affect the style of interaction. The same fully participatory conversation with three people does not occur with a group of one hundred. The individual adopts the proper role for that group, but the group defines what type of interaction is possible or appropriate.
Perhaps the broadest indicator for group interaction is group size. Of interest for a long time, Hare catalogs works on group size from the 1900’s to the 1970’s (Hare, 1981). In this work, the main finding demonstrates groups change with size. For example, a leader is more obvious and known as the groups get larger, while at the smaller scale they are less likely to emerge. At the same time, group members are more likely satisfied with the decisions of a small group — presumably because they are able to impact the direction. Early in group studies, researchers recognized group size as a strong influence on how a group keeps order amongst individual in the course of all interactions (Simmel, 1902). Different structures of communication are inherit as a group grows: business partners or a married couple do not communicate in the same way a military or aristocratic body might. In collection, these works recognize a hierarchy of group sizes that define ambiguous points where conversation changes. For the purpose of study, groups are often considered as the dyad, the triad, the few, groups of 10, and large groups (Simmel, 1902; Lindsay, 1972). Some researchers combine groups or set the numbers for large groups at slightly different levels. However, the distinction generally remains.

In casual conversation, where there are no pre-established leaders, groups are often limited to around four or five individuals. At more than four, a group has a tendency to splinter into more conversations due to constraints of distance and ambient noise (Dunbar et al., 1995). The social environment forces this equilibrium due to our innate limits of separating voices from each other and background noise. Additionally, larger groups require individuals to be less active in participation and encourage conversation with one’s neighbors.

Groups define their conversations by establishing a shared understanding or common ground. Individuals in the group contribute to this common understanding as conversation progresses, debating any contentious points (Branigan, 2006). Once established by some subset of group members, the common ground becomes a part of the conversation accepted by all unless a participant publicly indicates disagreement. Silence constitutes agreement unless a public record is necessary, though people often shake their head or offer short verbal acknowledgment as clear signals (Krauss et al., 1977). Conversation is a collaborative build up of common ground that accumulates as the conversation advances (Clark, 1992; Wilkes-Gibbs and Clark, 1992). Discussion and questioning gradually refines the established common ground. Larger lecture based discussion works in much the same way, with the lecturer establishing most of the common ground and the expectation of student participation to question when necessary. If there are no questions, a teacher is either left to assume the class has established a common ground or must purposefully probe for understanding.

In a face-to-face context, the normal patterns of blocking and evaluation anxiety can relegate individuals to silence as a group gets larger (Diehl and Stroebe, 1987). In the interest of establishing a common
ground, groups squelch dissent through majority. Larger groups have a difficult time establishing full unanimity, so individuals defer to the perceived majority in the interest of achieving agreement. A person’s anxiety of evaluation regulates dissent into silent agreement. As US President Abraham Lincoln said, “It is better to remain silent and be thought a fool than to speak out and remove all doubt.” While a person’s image could benefit by adhering to this advice, the net result encourages silence and reliance on the rest of the group (Jones and Gerig, 1994). The pressure to conform is less prevalent in anonymous and asynchronous interactions than in face-to-face or group interaction (Baltes et al., 2002). Employees in an organization withhold disagreements from employers in face-to-face meetings that they will divulge via email (Sproull and Kiesler, 1991; Markus, 1994).

The coordination of turns is imperative to any conversation. Turns in conversation allow multiple participants to coordinate conversation in their group. In casual conversation this control can be organic and locally organized, though a debate or large lecture might require a more formalized structure (Sacks et al., 1974). The turns themselves are a constant negotiation between speakers. A speaker ends his turn by giving a yielding signal which others are free to act upon (Duncan, 1972). These signals vary from ending a sentence, to prompting with “you know,” to termination of hand gestures, to dropping in pitch, adjusting eye gaze, among others (Duncan, 1972; Kendon, 1967). Once a speaker yields with a signal, others may feel inclined to speak up resulting in overlapping speech until individuals yield to a single speaker (Schegloff, 2000). Overlapping speech without such a signal is an interruption. Increasingly, empirical researchers apply technology to automatically attempt to observe, model, and detect the nuances of conversation activity (Gatica-Perez, 2009).

My work shows how the use of social mirrors impacts group dynamics. With each project presented, the new social cues push groups toward a new self imposed equilibrium. Most prevalently, each social mirror that demonstrated the relative amount of speech in conversation pressured groups toward a more balanced conversation (Chapters 4, 5, 7). Talkative people don’t want to overtly dominate conversation and quiet individuals speak up more often to participate. At the same time explicit anonymous noted as a part of the visualization encourages participants to assert themselves, direct conversation, and leads to a higher satisfaction with the results (Chapter 5).

**Group Support Systems**

The use of group support tools have developed to improve efficiency and streamline decision making. Way to encourage and better the decision making process through technology, these systems focused on the efficiency of group agreement and information passing. Grudin argues this, and the field of
Computer Supported Cooperative Work, developed out of the early history of “office automation” (Grudin, 1994). Thus group support also includes a number of now familiar technologies like email and file sharing, but also systems to engage large audiences simultaneously.

As with group dynamics, the design of a group support system is highly dependent on group size, proximity and task (DeSanctis and Gallupe, 1987). Thus support might range from single room support (Biehl et al., 2007; Olson et al., 1993), distributed communications (Ishii and Kobayashi, 1992), to a computer mediated conference (Olson and Olson, 2006; Bly et al., 1993). These systems seek to identify areas that cause a loss in productivity and address it with a technology solution. Thus they might help to share information, encourage objective evaluation, prevent blocking, and discouraging free riding (Nunamaker et al., 1991).

Small groups in local spaces can benefit from augmented awareness in work activity. For example, publicly viewing a document helps cooperative groups remain focused on their tasks and goals when working to solve a shared problem (Olson et al., 1993). In this study, any group member could edit and make changes to alter the document. Though these groups were less likely to deviate or try more creative solutions, the greater coordination produced more complete and better answers. The continual awareness provided by computer based cues has also provided beneficial coordination to programmers (Biehl et al., 2007). Providing status cues to demonstrate where work is taking place in the code-base enables other members to predict and prevent potential conflicts.

These support systems also open the door to providing anonymous feedback to the group. In small groups of co-workers anonymous groups are as effective as non-anonymous groups (Connolly et al., 1990). In fact, they produced more original ideas than their non-anonymous counterparts, though the quality of the end result is no better. However, when using anonymity, the lack of repercussions for creating animosity can be a problem in some groups where there is not a shared goal that everyone is working towards (Donath, 1999).

Feedback in the Classroom

One area I specifically examine in group support is feedback in the classroom. Creating new channels of communication in the classroom is demonstrably beneficial to making the class more engaged and participatory. Not only are silent students harder to evaluate, but the lack of attempting to participate can lead these same students to avoid taking risks (Jones and Gerig, 1994). Chapter 5 presents my work in small group feedback as well as classroom engagement.
Educators have applied group support to mitigate the increasing size of classrooms. A common interface, Audience Response Systems provide multiple choice and true/false responses (Fitch, 2004; Stowell and Nelson, 2007). These interfaces are most often used when the lecturer explicitly asks a multiple question of the audience. To be effective, the lecturer has the difficult task of anticipating key moments to query the audience and must specifically structure a lecture to accommodate this new question-answer format. Though each system varies, many include specialized hardware which are either purchased by or provided for each student (Stowell and Nelson, 2007; Kay and LeSage, 2009). In the worst cases, when a lecturer does not incorporate interesting interaction into the lecture, the Audience Response Systems become automated attendance and quiz systems, which students grow to resent (Kay and LeSage, 2009).

Other feedback modalities such as text based systems provide opportunities for students to engage with each other (McCarthy and boyd, 2005; Yardi, 2007). Studies of dedicated course chat-rooms show students will chat about the lecture’s content to help explain concepts to confused classmates (Yardi, 2007). In addition to making help available, this style of active learning helps students communicate concepts to peers for a deeper understanding of the material. However, the chat rooms also encourage unrelated discussions, and potentially draw students away from the lecture. Outside of the classroom, text based systems similarly open public dialog into shared events through IRC, instant messaging, Twitter, and Facebook (McCarthy and boyd, 2005; Ebner and Reinhardt, 2009; Shamma et al., 2009). Though all of these side-channels can contribute to audience discussion, they often leave the speaker out of the loop during the event. In some cases these systems are a part of the talk. The text based back-channel allows a wider audience to participate in the questioning of a guest speaker (Harry et al., 2009). It builds a set of questions to ask at the end, though it can still requires some moderation.

Systems based on annotating slides provide permanent context. Viewers annotate by using flags to indicate confusion as in (VanDeGrift et al., 2002). Messages like “Explain” and “Question” annotate the slides during class time, but the viewer cannot contextualize the message with additional information. Writing systems, such as Classroom Presenter, benefit from writing as input while including the instructor in the interaction (Anderson et al., 2007). A tablet PC system, it allows students to mark directly on the current slide with a stylus. Others in the system see the shared contributions, questions, and answers on their own screens. Instructors can set up slides that encourage students to answer questions as an interactive group discussion. This method enables a broad sampling of student understanding and encourages active participation with the material.
2.4 Visual Conversation

Computer mediated communication research strives to improve and explore communication interfaces. Particularly in video based chat, visual signals often replicate face-to-face interaction. These signals include body language, facial expressions, eye contact, and gestures — feedback that is desirable but often lacking in traditional remote conversation [Bly et al., 1993; Ishii and Kobayashi, 1992; Olson and Olson, 2000]. The best are often expensive and require special setup in order to use them effectively. As a result, remote visualizations of conversation often construct abstract visual cues based upon the text or audio that passes between users as a means to convey additional information, as in Figure 2.4 (Donath et al., 2000; Erickson and Kellogg, 2000; Kim et al., 2008; Leshed et al., 2009).

Though Computer Mediated Communication has sought to imbue remote conversation with meaningful feedback, Hollan and Stornetta argue the goal is not to recreate face-to-face environments, but to provide an interaction that is in some way better than face-to-face (Hollan and Stornetta, 1992). They argue that any feedback that truly improves upon the conversation experience must be beneficial in the collocated environment as well — else people would be apt to use such cues instead of face-to-face.

![Figure 2.4: Visiphone in (Donath et al., 2000) connects two remote locations with a constant connection. The audio signals produce a visual stream of circles one side in blue one side in orange.](image)

Applying social visualization based feedback in face-to-face conversation influences group dynamics by encouraging and discouraging participation [Bachour et al., 2008; DiMicco et al., 2004; Kim et al., 2008; Sturm et al., 2007; Kulyk et al., 2006; Streng et al., 2009]. DiMicco et al. first examined visualization as a way to encourage idea generation through discussion, as a way to get a wider variety of people to participate (DiMicco et al., 2004). For their experiment, researchers provided information for participants to discuss and come to a group decision. The visualization presented the contributions
of each speaker as a histogram visualization on a nearby wall (Figure 2.5). In each conversation, some information was selectively shared with only one participant; the best solution could only come out if everyone shared their private information. While the study demonstrated weak support for increasing information sharing, the visualization had a greater effect on the levels of individual contribution to conversation. Though balanced conversation is by no means the desired result for every conversation, it was the target in this manufactured conversation. All members of the group were equals in the task, and the experimenters provided each with different information to share. The explicit labels of over-participating and under-participating in the visualization highlight the desire for a balanced conversation. The same researchers continued studying the effects of real-time and post-meeting visualizations to determine the best form and support for a group meeting (DiMicco and Bender 2007).

In the works that followed (Figure 2.6), balanced conversation was consistently seen as a result in spite of different settings and visualization styles (Bachour et al., 2008; Kim et al., 2008; Sturm et al., 2007; Kulyk et al., 2006; Streng et al., 2009). Even when there was not an explicit indication to divide the conversation time more equitably, talkative individuals yielded more often to the quiet participants. Research in the area attempts to integrate these interfaces directly into the conversation setting (Table 2.1) through the use of wall displays, tabletops, or mobile devices (Streng et al., 2009; Kim et al., 2008). conversational setting (Table 2.1). Most presented the feedback in a diagram, though one group found an ambient display in the form of a mural was most effective for their purposes (Streng et al.)

**Figure 2.5:** Work in (DiMicco et al., 2004) demonstrates the use of a wall projection and histogram to demonstrate participant contribution. Labels of “under-participating,” “participating,” or “over-participating” characterize each individual.
An Interactive Table for Supporting Participation Balance in Face-to-Face Collaborative Learning

Khaled Bachour, ... and

try to elicit these types of interactions. As seen in Figure 2, the paradigm breaks down the complex question under...

Metaphor or Diagram? Comparing Different Representations for Group Mirrors

Sara Streng 1, Karsten Stegmann2, ... requires prior specific permission and/or a fee.

© '() $%%& *+,-: &./#0#1%22/# #34%&40 ...50%6%%

Additional copies are available at the ACM Digital Library (McCarthy and Meidel, 1999)

Detailed account of the user study conducted to evaluate learning gains.
or interactions that are shown to be predictive of positive participation on a shared display. We embed this display a group that do not participate in the group process, Lower learning outcomes are observed for members of being in the background of the collaboration process. We embed this display in a group that do not participate in the group process, Lower learning outcomes are observed for members of...

Additional copies are available at the ACM Digital Library (McCarthy and Meidel, 1999)

2M 1I

Figure 2.6: Research in conversation visualization includes the exploration of tabletop displays, wall displays, and mobile displays. Though the visualizations differ, commonly there is a tendency to balance conversation amongst all participants.

Differing from my own work, all of these visualizations focused on an aggregated summary of participation rather than a continuous depiction of conversational activity. Work with mobile devices also began incorporating additional features including body movement and proximity [Kim et al. 2008]. Their work further demonstrated the use of group dynamic feedback in remote scenarios has a similar effect. Anecdotal evidence from these groups suggests the details of the visualization had an effect; however, my own experiment in Chapter 7 indicates that the visualization details play only a small role in this change.

Text-based conversation dynamics have not escaped the use of visualization either. These projects show that subtle interface changes affect group interaction in small conversations [Leshed et al., 2009] and auditoriums (Harry et al., 2009). Leshed et al. demonstrated that highlighting the sentiment of messages and words in custom chat-room fosters a more positive environment (Figure 2.7) [Leshed et al., 2009]. As a group chats, highlighting agreement and negativity in messages encourages more agreement amongst the group and fewer negative words. At the same time, the general awareness
Table 2.1: These works demonstrate collocated conversation visualizations focused on changing group dynamics.

of contribution, just like the visualizations of face-to-face conversations, demonstrated the desire to meet a self imposed social protocol is not limited to one domain. The public awareness and resulting accountability created influences the way people interact.

2.5 Ambient Displays and Augmented Spaces

A key goal for HCI has been the creation of integrated experiences that blend interaction between the physical world and the digital world. Ambient displays and augmented devices bridge the divide in creating a seamless interaction with our technology or data. Augmented spaces tie information displays to interactive devices in our normal spaces away from the traditional desktop or laptop. This allows a user to interact with data in a setting, when properly designed, that is a natural fit for the data or presented in a more useful location. Ambient displays go one step further by integrating
information into an aesthetic display that people ignore most of the time. This allows a person to acquire data when needed, but not draw attention from whatever the current task may be.

Augmented spaces imbue generally static objects with a more dynamic operation. The two main approaches at this have been to create a new interface in a convenient location (i.e. where the interaction is most appropriate) or integrate physical components into the interface. As an example of the former the Everywhere Display projects an interactive interface anywhere there is a flat surface to guide customers through a store or patrons in a museum (Pingali et al., 2003). The latter approach adopts physical components to manipulate and interact with a display. These tangible bits enable the direct exploration of architectural designs, mixed reality games, and merging information to physical objects (Underkoffler and Ishii, 1999, Ishii and Ullmer, 1997, Kakehi et al., 2005). Tangible devices are particularly prone to use tabletop displays as they offer a flat projection surface and a natural place to set objects.

The tabletop has become a popular setting to explore collocated interaction. The table resides in a centralized space around which people converse, work, eat, play, etc. For this reason, it’s a common choice for augmenting – though it also provides a unique set of problems to deal with such as orientation of content, multiuser modal input, multiuser mouse input, input precision, and territoriality of the physical space (Shen et al., 2006). Researchers explored this domain using capacitance to track users (Dietz and Leigh, 2001), auto orienting content (Shen et al., 2004, Kakehi et al., 2005), and studying the patterns of sharing at a table interface (Scott et al., 2004). These works focus on understanding how people want to use tabletop interfaces, how they share information in a publicly shared space, and how they compensate for the new restrictions of the domain. However there is no definitive device or configuration to study as real world applications for tabletop interfaces are not established in everyday environments.
The ambient device extends from the ideas of augmented spaces. These displays generally offer a quick impression of information to the viewer, though most of the time the interface is an aesthetic background display that the viewer ignores. As an example, the AmbientRoom presents a space that might incorporate abstract imagery like a wave pattern projected on the ceiling or a quiet background noise to indicate status information. Unlike traditional interfaces, the ambient device often associates custom mappings between the data and the representation. When incorporated into a space, those familiar with the mappings will have access to information that guests will not perceive as they do not know the mapping. As ambient devices have a tendency to be background devices, their design appropriates aesthetic qualities as an incentive to leave it on when not in use.

**Reflection**

Reflective computing invites the individual to think about their interaction, their daily routine, their family, etc. and to bring out thought on unconscious values that we hold but do not consider. These principles push both the designers and users to challenge the assumptions of technology use by stressing critical thought or encourage deeper thought in a museum. Interpretations should be flexible, feedback should be dynamic, and technology should probe new areas. In the light of new feedback, the viewer is free to evaluate it according to their current context. As an example, the Home Horoscope promotes the general well-being of a household by encouraging a re-evaluation of routine. The technology took on a supporting role; sensors throughout the house monitored activity levels to infer the general well-being of the household throughout the day. After analyzing observed activity in a home an ambiguously worded horoscope encourages introspection into current life. The fundamental principles of Reflective Design encourages some amount of ambiguity and self evaluation. Ultimately the systems defer any judgment of values to the viewer.

The use of visualization has been an effective way to encourage other changes to personal behavior by just reminding people of their behavior. For example, encouraging lowered energy consumption is a common domain that has seen the use of visualization make an impact. Raising personal awareness of energy use impacts individuals with a desire to change with regards to carbon consumption or comparing one's own energy to neighbors. In another work, indirect monitoring of financial transactions demonstrates environmental impact. From these sorts of actions, one can detect and visualize the impact of the
products one buys or the impact of one's traveling patterns (Schwarz et al., 2009). In an expanded form the same interface could be a reflection on finances as well. The best way to motivate people to impact the environment does not have a set form. The individual's information can go to all others or individuals. Each interface has its purpose and handles aggregation and privacy in its own way. As with any visualization, testing is necessary to get the best results for any domain.

In addition to environmental awareness, reflection interfaces influences personal health by directly monitoring of personal habits. Even before introducing technology, fitness has used social connections as a way to encourage perseverance and reaching the next goal. Consolvo, et. al present a similar motivation by automatically capturing one's exercise information and providing a display to indicate how successfully a person meets their exercise goals (Consolvo et al., 2009). At a glance, the display shows the activity for one week worth of goals (set by the user). A detailed view is also available as a log.

Social mirrors serve to augment the understanding of how others perceive us. We naturally develop a sense of others throughout life based on what we observe. Our mirrors augment this ability by showing behavior that goes unobserved or unrecognized. By understanding one's own interaction, a person better infers the thoughts, feelings, and desires of others (de Waal, 1996; Meltzoff, 1990). With a more diverse set of cues, the social mirror aides in the development of social adeptness.
I have collectively described my work under the theoretical construct of Social Mirrors to describe interfaces that encourage self evaluation and self initiated moderation. The interfaces I describe are analogous to the mirrors we use on a daily basis. In this thesis, social mirrors are real-time depictions of social interaction via visualizations integrated into the surrounding environment. The visualization is the metaphorical mirror, it offers a viewpoint from which to judge oneself. The reflection is based on what aspects of conversation the mirror can capture and redisplay back into the conversation. Typically, these mirrors specifically show traces of social dynamics that would otherwise remain unnoticed in conversation. Once aware of these dynamics, a viewer can choose to alter his or her actions in order to better present themselves to the rest of the group. Consider the use of traditional mirrors:

*The morning of a presentation, you wake up and think through your plan for the day. After a quick morning shower, you brush your teeth and run a comb through your hair. Mentally reviewing your presentation, you check your face, skin, hair, and then realize your part isn't even; you re-comb it.*

*In the bedroom, you grab the clothes hanging on the door; a dress shirt, jacket, and slacks you set out last night, they look stylish enough to indicate you care about your appearance, but not overly pretentious. Looking in the full mirror on the door, fit looks good, colors are fine, and those trips to the gym have paid off. Then you notice the shoes: scuff marks. No time to polish, you switch to another pair. You decide to leave the jacket at home, as it looks too formal. You're off to work.*

*You walk up to the glass door and see the wind disheveled your hair a bit, but easily fixable. You head upstairs, confident your appearance fits the part.*

Our appearance to others is based on so much more than our physical looks. The social mirror builds upon the self-evaluation affordances of the traditional mirror to allow a person to see how they appear to others. While styling our hair or adjusting our clothes, we evaluate our personal appearance as seen
in the reflection. We continue to style and adjust until that reflection matches an image we wish to show the world. The social mirror lets us view and change our actions and our language based on our appearance to others (Chaika, 1994; Cleveland, 1888; Goffman, 1963). The social mirrors in this thesis use abstract visual imagery stemming from social visualization literature to convey this insight immediately and quickly.

3.1 Foundations

The foundations of social mirrors are widespread and varied. The last chapter highlighted these areas including signaling theory, group awareness, visualization, and reflective design. Combined, these areas are meant to facilitate reflective thought in order to better adapt to the necessary, or desired, roles in conversation. When applied to the setting of the classroom, therapy, or meetings, a person can use that visualization to better fit into the social environment, to assert themselves, or to initiate conversation.

What defines a social mirror?

Unfortunately, I cannot claim to be the first to use the term Social Mirror. More so, I am not the first to use it in the sense of self-evaluation. The term has seen a wide use in a number of different areas, though it does not seem to have become common parlance in any of them.

However irregularly the term may have been used, each tends to emphasize the context of understanding something from the viewpoint of society or a large cultural group. As examples, the comparisons made at a high school reunion was described as a social mirror (Vinitsky-Seroussi, 1998). Each person could judge their own success by comparing the relative success of others raised in similar environments. A book on etiquette used the term social mirror to describe the awareness of proper protocol to interact and remain in high society (Cleveland, 1888). The Planet of the Apes has been described as a social mirror for showing audiences an image of what our society can be like (Nichols, 1998). As a final example, children’s drawings have been described as social mirrors because they depict and incorporate gender stereotypes that are perceived though not explicitly taught (Margolin, 1994). As in my own definition, the social mirror reveals aspects of interaction and culture that may otherwise go unnoticed.

I define a social mirror as a live visualization generated from traces of interaction and presented directly to the same people with the intention to provide new insights about their social personae.
Many of the types of cues I seek to demonstrate are temporal in nature: turn-taking, interruption, conversation domination, etc. As a result, people will forget details of conversation and interaction minutes, even seconds, after they occur (Bernard et al., 1984). There is little chance to perceive how one appears to others. The social mirror provides that insight by reflecting these cues back to the participants. The individual can focus on the salient aspects and change their social appearance as they see fit.

As I aim to encourage introspection into one’s own social signals for the purpose of initiating action, the social mirror breaks into three component stages:

1. **The social mirror captures social signals.** A person constantly sends off signals and cues when communicating. The social mirror must reproduce these interaction patterns visually as they occur.

2. **Viewers can observe the constant stream of signals.** Not only must the signals be perceivable, but it must be interpretable. In creating new cues, the mapping from visualization to action must be clear: a reflection has less value if it is inaccurate.

3. **Social mirrors encourage self-evaluation.** The signals we see in the social mirrors are signals we show to others, perceive in others, but rarely see in ourselves. The new perspective, akin to a third person viewpoint, allows comparisons directly with others and through their eyes. This evaluation allows the individual to either accept their current interaction or initiate change.

Using a social mirror, a person has the opportunity to better understand their personal interaction as well as the group level dynamics. Social mirrors allow the individual to focus on the salient aspects and change their social appearance as they see fit. The social mirror offers a external feedback point to aid in the navigation of social environments and improve the understanding of the group.

### 3.2 Hall of Mirrors

This section takes a brief look at the projects I will present as a part of this dissertation. As an introduction to these projects, I demonstrate the basics of each system: what they capture, what they show, what they demonstrate, and what value it provides. Later chapters provide further details of the experimental setup and provides the details that justify my conclusions.
I wish to emphasize that social mirrors are not limited to the styles presented in this dissertation. I see my own mirrors as a specialized subset that alludes to a larger domain. Each project is the result of design decisions for visual appeal, desired functionality, and environmental integration. I based the design of each on related literature, initial prototypes, critical feedback, and my own intuition. However, the overarching ideas of social mirrors permeate them all.

Each of these projects involves unique constraints and design decisions. The visualization must be tailored to the environment and be visible to all participants. I adopted a policy of minimalist design to provide the most information with the least distraction. It was as much a decision to decide what to show as it was to decide what not to show.

**Conversation Clock**

The initial Conversation Clock showed how individuals will balance their contribution in conversation. The visualization (Figure 3.1) demonstrates patterns of interaction: conversational dominance, turn taking, non-participation, volume, and the rhythm of conversation. It presents a persistent depiction of past speech by showing the continually sampled conversation. The ephemeral moments of conversation remain long enough to examine. The structure of the Conversation Clock ensures a glance to the table will be informative while keeping details accessible to the viewer. With no other instructions other than to talk, the visualization encourages the talkative participants to speak less, and the quiet to speak more. The conversations, which had no preset hierarchy, created a more equitable distribution.

![Figure 3.1](image)

**Figure 3.1:** The image of the Conversation Clock presents a circular representation of conversation. Each bar a moment of time, the captured aural history defines the concentric rings. The visualization highlights characteristics such as domination, silence, and turn-taking.
The Conversation Clock does not actively or explicitly encourage a balanced conversation; this goal emerged from the groups themselves. Participants indicated they were more aware of how they appeared to others. The visualization provides an extension of signals and cues that reflect on the individual, an extension of Goffman's theory of self (Goffman, 1959; Goffman, 1963). Not wanting to appear dominant or uninvolved, participants adjusted their behavior based on the social mirror. The talkative spoke less in each turn; the quiet spoke more often. In agreement with the quantitative data, all groups reported the visualization pushed them toward a more balanced interaction.

Originally, the Conversation Clock study appeared as part of my master's thesis (Bergstrom, 2006). In this dissertation, it appears as Chapter 4 to provide adequate context and background for the social mirror's exploration. This first study characterized how a visualization could be useful while embedded in the conversation environment. It directly lead to additional questions that I explored in later work. Figure 3.2 characterizes the three main branches of inquiries that developed following the initial Conversation Clock study.

![Figure 3.2: The initial study with the Conversation Clock began explorations into the use of feedback, emphasis, and content incorporated into the social mirror.](image-url)
Feedback

Having seen that the visualization created a new awareness of interaction, I explored the use of the interface for explicit feedback from the crowd. In Conversation Votes and Fragmented Social Mirror, I found it to be an effective communication mechanism for anonymous interaction.

Emphasis

The visualization's feedback affected people. They reported greater awareness of their own interaction, but were they really following the visualization's depictions? Applying distortion to the Conversation Clock I investigated the effects when the visualization was not an accurate depiction.

Content

When dealing with conversation, content is always a question. Conversation Clusters was a prototype interface that captured the continually changing conversational topics. A complex project, it demonstrated the feasibility of capturing real-time topics from conversation and incorporated them into visualizations.

Feedback: Conversation Votes

Having shown visualization can communicate passive feedback into conversation, I developed Conversation Votes as a means for people to direct their personal interpretations of conversation anonymously. Less personal forms of feedback encourage direct conversation across artificial barriers: email had this effect on company hierarchies (Sproull and Kiesler, 1991). The new visualization allowed viewers to indicate a binary indicator of agreement/disagreement.

While the original Conversation Clock visualization affected how people contributed to conversation, there was no distinction between eloquent monologues and random babble. The new design allowed participants to indicate agreement and disagreement directly in the visualization by pressing inconspicuous buttons. The feedback was publicly visible, but anonymous within the small group.

The feedback channel also offered an outlet to anyone not able to speak. In face to face group situations, social pressure regulate the less outspoken to silence. The result is fewer voices, fewer ideas, and group-think. Mediated channels, such as email, allow more people to join in discussion and offer their opinions. Conversation Votes allowed people to express a basic opinion (agreement/disagreement) anonymously to avoid any evaluation anxiety. As only a relative few can speak simultaneously, the voting feedback enables a parallel communication channel.
Anonymous feedback proved to be an enabling channel for participants. The study consisted of debates between group members. People who most used the channel felt the opinions discussed were not adequately representative of their own opinions. In conditions with the visualization and incorporated feedback, people were able to better assert themselves by speaking more in the conversation. They spoke more, and felt better about the debate.

**Feedback: Fragmented Social Mirror**

Building from the small group, the ideas of the Conversation Votes influenced the Fragmented Social Mirror. This social mirror applied anonymous feedback to a larger group - a lecture based classroom. Each student has a personal interface to ask questions and provide clarification during a lecture. During a small experiment, students asked more questions. Perhaps more importantly, they initiated questions to the professor, wherein previously they did not.

The social mirror in this case was fragmented: the interface was distributed throughout the class and the purpose of the mirror differed for roles of student and lecturer. The reflection in this case was a sense of the students understanding as well as questions for the lecturer, but also it reflected to the students that others had questions as a way to encourage further questioning.

*Chapter 5 presents further details on Conversation Votes. The lessons learned from Conversation Votes are also directly influence the feedback in the Fragmented Mirror discussed in Chapter 6.*

**Emphasis: Conversation Depiction versus Public Accountability**

The Conversation Clock and the Conversation Votes visualizations both pushed people toward a more balanced conversation. The question still remained: why do people change their interaction? Do they react to their portrayal directly or to a sense of public accountability? If people react to their depictions, a visualization could encourage or discourage a specified individual's participation by actively moderating conversation. If they react only to a sense of public accountability, setting explicit public goals becomes more important to shape conversation.

Using an altered version of the Conversation Clock, I experimented to investigate the notion that the visualization is directly influencing conversation. In this study, the Conversation Clock visualization purposefully misrepresents speaker contributions; an individual would appear to speak much more than they actually did. Using four different strategies, participants were over and under represented in
conversation. The study showed that people are aware of their depiction in the visualization, but that alone doesn’t directly influence interaction.

The implications of this work apply directly to ambient awareness and visual displays. Consider using shared visualizations to demonstrate energy consumption or daily exercise. The knowledge of public monitoring would likely produce some effect due to a sense of accountability. However, a comparative data visualization alone could only go so far to influence behavior.

**Content: Conversation Clusters**

A central question in dealing with conversation, how can you capture and represent the content? I began to explore this area as a way to create an archival tool and meaningful artifact. Conversation Clusters incorporated speech recognition and clustering algorithms to generate models of conversation in real time. It uses these models to extract and display thematic development of conversation over time.

The initial prototype demonstrated the viability of extracting meaningful topic words directly from conversation using Explicit Semantic Analysis (ESA) as trained on a snapshot of Wikipedia. Each article in Wikipedia represented a collection of words to identify the topic (article). My representation, built on top of ESA, examined the most salient words that identified topics, and combined them into related clusters. The interface presented an artistic representation of current topics on the table at all times (shown above).

Conversation Clusters enabled a human-computer dialog to work towards better solutions. The algorithm was not perfect, though it was able to detect meaningful topics for most conversations. Unrelated words or clusters did appear were editable via touch sensors. Changes made on the table refined the underlying clustering system.

Conversation Clusters has room to grow in my future work. I have shown how to detect and archive meaningful topics visually during conversation, in spite of complications due to untrained voice recognition. I have yet to fully study the use of thematic archival of meetings over the long term with regards to memory aids, retrieval, and summarization. I consider this to be a project to pursue in future work.

*Chapter 8 presents further details about Conversation Clusters.*
CHAPTER 4

Conversation Clock:
Raw Feedback as a Social Mirror

The Conversation Clock was the first exploration of the connection between real-time visualization and its effects on conversation. This first experiment focused solely on presenting "just the facts" of conversation. The Conversation Clock shows the details of conversation as they occurred and notes the speaker: it does not provide any interpretation of desired speech. These details include the turns taken, simultaneous speech, and silent pauses; but we also wanted to capture the larger picture to contrast conversational dominance with the lack of participation. The Conversation Clock provides a real-time view of conversation dynamics while leaving the interpretation of context to the viewers.

The Conversation Clock experiment shows that people are aware of these artificial cues and use them to evaluate their own participation. Using the Conversation Clock in a casual conversation, people adjusted their speech and turn lengths according to their visualized appearance. At a glance, the participants proved to be aware of their depiction in conversation and sought to change it. Talkative participants would back off and the quieter participants would assert themselves more often. These conversations demonstrated the Conversation Clock effectively nudging individuals to more equitable contributions in conversation.

This chapter\(^1\) details this first experiment and lessons. Once we learned how individuals paid attention to the visualization, we used this knowledge to enhance the communicative ability of this social channel in later extensions.

4.1 Design Choices

The Conversation Clock was the first social mirror developed as an exploratory interface. The design leveraged background theory and iterative design testing as described in the following paragraphs.

\(^1\)This chapter also appears in my master's thesis [Bergstrom, 2006]. It is the beginning of my work on social mirrors, presented here for completeness. This exploratory experiment tested the initial impact of visualization.
During face-to-face conversation eyes play a strong role in to indicate speaker and interest (Duncan, 1972; Kendon, 1967; Goffman, 1967). The design of the Conversation Clock encourages participants to examine the visualization to assess conversation. However, I purposefully sought to avoid visuals that might distract a person from conversation at inopportune times. The initial prototypes helped to determine what aspects were most informative or most distracting. Each design underwent informal testing to determine the most useful aspects and future designs. Since groups would sit at the table for brief conversations and view the visualization from different angles, it was important to keep the image non-oriented; it should be readable from any angle.

Figure 4.1: Here we see some of the initial design iterations. Each design underwent informal testing to determine the most useful aspects and future designs. Groups would sit at the table for brief conversations. The design was iterated weekly and refined into the Conversation Clock.

Participants indicated the heavily structured timeline helped to interpret the data, while excessive animation tended to distract speakers. Some of the earliest designs were eliminated for being too distracting. One in particular was described as reminding a viewer of a “flushing toilet” as the history spiraled towards the center (Figure 4.1). Others lacked structure and were seen as overly complicated and hard to read pie charts. We also dropped a number of features from designs, such as pitch, as they were considered hard to understand in our chosen visualizations.
The earliest designs of the Conversation Clock was focused on making the table a self contained object (Figure 4.2). Initially, the microphones were directly mounted in the table itself, in order to allow a person to just sit down and start talking. Eventually, the thinking went, the microphones could just disappear. Unfortunately, the pilot tests demonstrated that it was far too easy to pick up on background table noise, finger tapping, and bumping the table. The design changed to use lapel microphones as a result. People were less “free” to move around, but it ensured a clear input signal.

Prior study of conversations shows there is a complicated interaction of gaze in conversation, but that there is a lot of time where gaze fall away from the audience (Kendon, 1990). Both speakers and listeners have a tendency to look away while still attending to conversation that we hope to use as opportunities to view the social mirror.

The design of the Conversation Clock ensures the visualization is centrally located and readily available to all viewers. The visualization evolves slowly over time, so as not to purposefully draw attention to itself. To avoid distracting the speakers, each individual sample only changes a small portion of the display, ensuring minimal attention lost (Simons et al., 2000). As human peripheral vision is motion sensitive, the Conversation Clock was design to be static as much as possible.

Finally, the Conversation Clock demonstrates a visual preference for recent interaction. As memory in conversation is ephemeral, this highlights the most recent events while still retaining a longer history of overarching patterns. In contrast, the visualization can demonstrate a detailed history of conversation. Instead, we focus primarily on the recent past as people will have forgotten significant details after
only a few minutes (Bernard et al., 1984). For this reason, the visualization compresses past minutes into a central summary.

4.2 The Conversation Clock

The Conversation Clock’s name refers to its circular design and semblance to a clock. Our initial designs converged on a circle to target small group interaction. With four individuals in a conversation, everyone can participate in a central discussion as side-conversations are less likely to occur. The circular design ensured all participants received a similar view of the visualization and minimized any orientation effects. A circular table paired with the visualization embeds the projection naturally in the physical space (Figure 4.3).

![Figure 4.3](image)

*Figure 4.3: The Conversation Clock demonstrates simple patterns of conversation based on microphone input.*

The Conversation Clock visualization provides an accurate depiction of conversation by listening in on the conversation. All people at the table wear a microphone that the social mirror then monitors. By continually sampling the microphones, we know when a person spoke and how loudly. The clock captures the immediate moment as a single bar on the table. The longest bar at any moment identifies the loudest speaker. For the purpose of later analysis in this and later chapters, we define a *lead* as a moment when a person is the loudest speaker, whether simultaneous or solo; a *turn* is a count of how
many times a person gains the lead; and \textit{turn length} measures the duration of each turn. Overlapping speech creates overlapping bars in the visualization; the smaller of which indicate the speakers that were quieter than the lead. Figure 4 presents common characteristics of conversation and how they appear in the Conversation Clock.

As time progresses, the visualization grows to include new interaction as it occurs (Figure 4.5). The outer circle always shows the most recent speech and adds new samples until a full minute elapses and completes a ring. Upon completion, the ring is compressed inward as a part of the conversation history. The rings build up in concentric circles to show the passage of minutes much like a tree's rings depict its history with the passage of years.

The visualization is a simple transcription of audio patterns that reveals ephemeral cues; however, this first interface refrains from presenting contextual cues. The viewer judges the visualization in their personal context. The Conversation Clock does not try to make sense of words, it simple presents an unbiased reflection of what it monitors. However, the context is necessary to interpret the visualization.

4.3 \textbf{User Study}

The first study of the Conversation Clock served as an exploratory tool. Prior to this work, DiMicco et al. showed that bar chart depictions of participation can affect the dynamics of groups (DiMicco et al., 2004). We anticipated similar results as the relative participation is still apparent in the visualization, though not as easily comparable. The Conversation Clock was an exploration of real-time feedback at a finer scale such that individual moments of interaction are visible. Specifically, we investigated:

\begin{itemize}
  \item How do people react to an immediate transcription and reflection of their group interaction?
  \item Can conversations progress naturally with such a visualization in place?
  \item Is the visualization distracting?
  \item Does the visualization seem to limit or encourage behavior?
\end{itemize}

These questions were intended to define a direction for future work rather than set a hypothesis. Aside from the DiMicco's et al. prior work showing a tendency to balance conversation, there was little to judge how people might perceive such a visualization projected directly into the conversation space. The Conversation Clock offers an exploratory look into the this space and informs the design of later studies.
**Turn Taking, Interruption**

Patterns between turns appear as a change between colors. In some cases, it appears as an immediate change from one color to another. In others, overlapping speech is seen either as an interruption or a part of natural conversation protocol. In the end, it’s up to the viewer to determine which it was.

**Agreement, Aural Backchannels**

Simple verbal backchannels such as “yeah” or “uh-huh” appear as short bursts inside longer dialog.

**Silence**

With the ever progressing visualization, periods of silence appear as empty space punctuated by small dots. A conversation with lots of silence appears spread out in-between each turn.

**Dominance**

Dominance in conversation can be seen as one color appears much more prevalent than the other. As history builds up, a full impression of domination can become apparent.

**Mimicry, Rhythm and Flow, Time Spans**

Other aspects that pertain to the overall conversation are also apparent. Mimicry, in that people tend to adopt a similar volume, during conversation becomes apparent. While the Overall rhythm and flow can be seen as a generally equally divided conversation or a lecture/leader based conversation based on the length of relative turns over time.

**Figure 4.4:** Above are examples of overlapping speech, silence, turns, and other characteristics shown in the Conversation Clock.
Figure 4.5: In the above images, we see how the Conversation Clock structures samples over time, clockwise around the circle. When a ring is filled, the visualization compresses to make room for the next minute.

Configuration

Four lapel microphones connect the participants to the Conversation Clock. Tags on each individual microphone indicated the color associated with that microphone. During setup, we inform participants of their color mapping. A DV-camcorder monitored interaction from a nearby tripod during all sessions. During conversation, this video recorded gestures and physical references for later analysis.

Methods

The user study took place in the HCI/Graphics lab in the UIUC Computer Science building over a one-month timeframe. Participants committed to meeting, as a group, for three 30-minute sessions with a week in between each session. However, if groups finished early we captured the natural break up of a session and did not force interaction in order to extend the time. The week-long break ensured participants were fresh for the next sessions.

We solicited participants as groups rather than individuals. Each member had to be familiar with all other members of the group and have a predefined purpose for meeting to define the theme for all three sessions. We chose to use familiar groups in order to ensure a more comfortable and natural interaction environment. We felt that using random groups over the course of three separate sessions would alter interaction patterns as individuals became more familiar with the other participants in their group. The focus of group conversation covered the topics:

- Weekly meetings for a class project  (3 groups)
- Conversation on old and current movies  (1 group)
Each group met for three sessions. The first session had no visualization present. A researcher observed and recorded the entire interaction for a baseline description of the group's interaction. In the second session, we introduced participants to the Conversation Clock. The final session, without the visualization, allowed us to see if the Conversation Clock had any residual effects. That is, since we expected the participants to be familiar with their image in the mirror from the second session, do they continue to feel pressure from that session.

In all sessions, participants wore microphones and underwent a calibration process. We gathered data regarding the speech patterns of individuals by extracting contribution information from the Conversation Clock logs. Though only projected for the second session, the Conversation Clock collected data for all three sessions in order to generate comparable data logs of each conversation for analysis. From these logs, we extracted the number of turns, length of turns, amount of speech, and other descriptive measures.

With the visualization active on the table, we informed participants of the color mapped to their microphone. Each participant had a card of the same color in front of them as an additional physical reminder. The groups took part in a short preliminary conversation to become accustomed to the visualization before the second session began. We instructed each participant to speak at least once to better understand how the visualization worked and to test our calibration process.

Within the visualization condition we used two settings for the visualization history. Each group would use the Conversation Clock with a visualization history of 5 minutes in a 10 minute session and a history of 20 minutes in a 20 minute session. To account for possible ordering effects, half of the groups were randomly chosen to receive the short history first. The researcher interrupted after 10 or 20 minutes (appropriately) to switch conditions. Between the sessions, participants filled out a brief survey as we logged the past data and reset the Conversation Clock.

4.4 Study Results

The Conversation Clock demonstrated a balancing effect of conversation of which even the participants were aware. The surveys following the sessions show that participants reported their interaction to
have been moderately altered by the table. Additionally, using a Wilcoxon Signed Rank Test, we found that the participants perceived a greater effect of the visualization with the 20 minute session with a longer history (Table 4.1).

<table>
<thead>
<tr>
<th>Measurement</th>
<th>10-minute</th>
<th>20-minute</th>
<th>Wilcoxon Z</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distracting</td>
<td>2.56</td>
<td>2.80</td>
<td>-1.364</td>
<td>$p &lt; .2$</td>
</tr>
<tr>
<td>Helpful</td>
<td>2.46</td>
<td>2.36</td>
<td>-0.966</td>
<td>$p &lt; .4$</td>
</tr>
<tr>
<td>Altered You</td>
<td>2.48</td>
<td>3.00</td>
<td>-2.871</td>
<td>$p &lt; .004$</td>
</tr>
<tr>
<td>Altered Others</td>
<td>2.65</td>
<td>3.02</td>
<td>-2.066</td>
<td>$p &lt; .04$</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>3.24</td>
<td>3.04</td>
<td>-1.475</td>
<td>$p &lt; .2$</td>
</tr>
</tbody>
</table>

Table 4.1: Results of the Likert Scale surveys in the Conversation Clock user study. Using a Wilcoxon Signed Rank Test, we show that participants reported a greater effect of the long history than the short history.

The level of satisfaction with the Conversation Clock varied greatly (ranging from 1 to 5) depending on the individual, group, and task. In aggregate, participants were neither particularly satisfied nor greatly dissatisfied. Participants reported the Conversation Clock to be somewhat distracting and somewhat helpful.

In addition to self-reported measures of interaction, coded video provided additional measures. Each session with the Conversation Clock was hand-coded to note gaze direction and gesture position. Gaze direction consisted of “at the Conversation Clock,” “toward other participants,” and “away from conversation,” while gesture position was “over the table/visualization” and “away from the table.”

<table>
<thead>
<tr>
<th>Metric</th>
<th>F value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Averting Gaze</td>
<td>$F_{3,3} = 15.56$</td>
<td>$p &lt; 0.03$</td>
</tr>
<tr>
<td>Table Gestures</td>
<td>$F_{3,3} = 5.98$</td>
<td>$p &lt; 0.09$</td>
</tr>
</tbody>
</table>

Table 4.2: For analysis, we hand-coded video taken from each sessions for gaze direction and gestures over the table. Using a repeated measures ANOVA, we found significant changes in the gaze direction of participants. A notable, but not significant, change in table gestures was also observed.

Video coding and a repeated measures ANOVA revealed participants significantly changing their gaze patterns ($F_{3,3} = 15.56$, $p < 0.03$) by averting their eyes more often from the conversation and to the Conversation Clock (Figure 4.6). A notable change in the table gestures appeared, though our sample could not show significance ($F_{3,3} = 5.98$, $p < 0.09$).

Further examining the change in gaze patterns induced by the Conversation Clock, we compared the 10-minute and 20-minute sessions via paired t-test ($t_{6} = 0.94$, $p < 0.38$). The test proved insignificant, indicating gaze patterns were not greatly affected by the length of visualized history.
Figure 4.6: The Conversation Clock increased the time spent looking away from others in conversation. Most of the looks to the Conversation Clock were in addition to the baseline set in the non-visualization sessions.

<table>
<thead>
<tr>
<th>Participation</th>
<th>Metric</th>
<th>F value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above Average</td>
<td>Leading</td>
<td>$F_{3,10} = 1.29$</td>
<td>$p &lt; 0.3$</td>
</tr>
<tr>
<td></td>
<td>Turns</td>
<td>$F_{3,8} = 1.19$</td>
<td>$p &lt; 0.4$</td>
</tr>
<tr>
<td></td>
<td>Length</td>
<td>$F_{3,5} = 9.22$</td>
<td>$p &lt; 0.02$</td>
</tr>
<tr>
<td>Below Average</td>
<td>Leading</td>
<td>$F_{3,7} = 0.52$</td>
<td>$p &lt; 0.7$</td>
</tr>
<tr>
<td></td>
<td>Turns</td>
<td>$F_{3,9} = 3.89$</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Length</td>
<td>$F_{3,12} = 3.32$</td>
<td>$p &lt; 0.06$</td>
</tr>
</tbody>
</table>

Table 4.3: Audio data, automatically logged and analyzed, was available for each participant. Expecting different changes to occur in each, we categorized participants as talkative or quiet based upon the initial session. We ran a repeated measures ANOVA and found talkative participants change the length of their turns while quiet participants changed the number of turns.

When examining the audio data, we split participants into two categories. Based upon the initial non-visualization session, we categorized participants as either talkative or quiet. As shown in DiMicco’s earlier work, we expected the different categorizations of individuals to alter their participation differently (DiMicco et al., 2004). Made strictly for analysis, the participants were unaware of their categorization.

Running a repeated measures ANOVA, we found significant alteration in some areas (Table 4.3). Specifically, we noted reductions in turn length of the talkative participants and the increase in the number of turns taken by quiet speakers. Having seen significance overall, we investigated the two visualization conditions. Once again, we found no indication of notable differences between the short and long history.

The Conversation Clock’s observed data was also sorted to examine how interaction changed over time. As no significant differences were observed between the 10 and 20 minute sessions, we became interested in how interaction changed with continued exposure. We see in Table 4.4 the most affected audio feature was the amount of time spent leading the conversation. This is true for both the talkative
and quiet categorizations.

<table>
<thead>
<tr>
<th>Participation</th>
<th>Metric</th>
<th>N</th>
<th>Condition 1</th>
<th>Condition 2</th>
<th>Paired t</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above Average</td>
<td>Leading</td>
<td>15</td>
<td>20.24</td>
<td>16.85</td>
<td>$t_{14} = 2.66$</td>
<td>$p &lt; 0.02$</td>
</tr>
<tr>
<td></td>
<td>Turns</td>
<td>12</td>
<td>1.79</td>
<td>1.67</td>
<td>$t_{11} = 0.839$</td>
<td>$p &lt; 0.4$</td>
</tr>
<tr>
<td></td>
<td>Length</td>
<td>11</td>
<td>14.87</td>
<td>12.21</td>
<td>$t_{10} = 2.16$</td>
<td>$p &lt; 0.06$</td>
</tr>
<tr>
<td>Below Average</td>
<td>Leading</td>
<td>8</td>
<td>6.86</td>
<td>11.20</td>
<td>$t_{7} = -3.56$</td>
<td>$p &lt; 0.01$</td>
</tr>
<tr>
<td></td>
<td>Turns</td>
<td>11</td>
<td>0.98</td>
<td>1.19</td>
<td>$t_{10} = -2.06$</td>
<td>$p &lt; .07$</td>
</tr>
<tr>
<td></td>
<td>Length</td>
<td>12</td>
<td>7.69</td>
<td>9.30</td>
<td>$t_{11} = -1.13$</td>
<td>$p &lt; .3$</td>
</tr>
</tbody>
</table>

Table 4.4: Re-sorting to indicate the first condition seen and the second condition seen, it becomes apparent there were novelty effects. Participants in both the talkative and quiet participation categories altered the amount of time spent leading. Examining the other notable statistics, the talkative category seems to have altered the length of their turns while the quiet participants seemed to change the number of turns to accomplish this feat.

![Figure 4.7](image.png)

Data coded from the video is also emphasized by a chronological examination. Participants altered both the gaze patterns ($F_{3,4} = 24.5, p < 0.005$) and gestures to the table ($F_{3,4} = 10.20, p < 0.02$), according to a repeated measures ANOVA. Examination of the actual data shown the Conversation Clock still increasing the number of gazes averted from conversation and objects referenced in conversation and the decrease in the number of gestures (Figures 4.6, 4.7).

Qualitative

Participants perceived the first session to be a normal interaction. As we purposely chose to use groups that were already formed and familiar, participants were in a familiar scenario. We consider the lab environment to be a small change, as three of the seven groups we studied do not meet at a regular location. Some participants made comments on the setting, but they were no more critical than:

"After a while it was [normal], once we got on topic... That and we had to sit at a table."
Observing the interaction of the groups during the sessions, we noticed that these initial sessions had skewed participation levels. This included three groups that had a combined four participants with less than 10% of contributed speech in conversation. Only one group reported an unbalanced conversation. Another stated it was typical to what they had seen in past meetings.

The second session displayed our social mirror, the Conversation Clock. Participants seemed more aware of their own and others’ participation.

“I realized that I could monitor my speech patterns by watching the colors. It was interesting to train myself not to say ‘ummm’ as much or pause.”

“It’s easy to judge who is driving conversation.”

“I was trying to look at the circle to see whether we were balanced.”

“It’s more salient, what’s going on.”

During one group’s discussion, a collaborative decision on what movies to watch at a weekly meeting, participants felt there should be a balance between themselves. However upon seeing the visualization they realized how heavily dominated the conversation was:

“I noticed when you’re the one talking, you want to stop. But if you’re mid-topic you couldn’t stop, because you had to finish your topic. But as soon as you finished your topic, you’d shut up.”

When we asked the same group if their interaction was a balanced conversation the heaviest contributor (about 40% of the leads and 40% of the turns) responded, “Not after looking at the clock.” This group also indicated an expectation of equal participation in this group, whereas a similarly ‘unbalanced’ viewed the contributions differently,

“Project managers communicate more than testers or developers, thus it is reasonable for it to be unbalanced.”

Another participant making a heavy contribution to conversation commented on the completed ring as a milestone, noting:

“By the time the ring had returned to the starting point I was like ‘That’s too much’ and a stopped talking because I feel like I completed that ring. That’s one unit that’s mine.”

Whereas participants who had been silent for a while could be prompted or compelled to speak by the Conversation Clock.
“Based upon more of the history after a while, when I did glance at the clock and then it like
reminded me to prompt [Blue] or direct questions to him.”

“It became all red, should green or yellow speak next?” - Yellow

Very little conversation during the sessions directly referenced the Conversation Clock. Over the
course of conversation a group would generally only mention the Conversation Clock one time if at all. During the experiment, the focus was on topic for their group. However, during the break between
conditions and after the experiment was over, participants focused direct attention to the clock. Some
just spoke to watch their contribution, others would yell and try to make their own color dominate, another group began to whistle. Groups reported it was “fun,” “interesting,” and “amusing.”

The Conversation Clock encountered criticism in groups with a task based purpose or a more defined
schedule. Groups with more structure and predefined interaction had less flexibility. One participant
explained his lack of contribution,

"Due to our group roles each member was expected to talk about a certain project anyway,
so I had little input on something I didn't know."

Some participants also found the visualization distracted them from the meeting at hand.

“I found it a bit distracting, I don't like things in my visual space.”

Though Conversation Clock was not visible in the third session, participants reported thinking about
it during their interaction. One participant stated:

"The first [session] I wasn't aware of it ... the second [session] I looked at the clock and tried
to make it balanced. Now, I don't have a measure ... but I still tried to use the impression
from last time. Because he always talked more [last time], I tried to consciously talk more
[this time]."

Another participant who had been the heaviest participant in his group during session two commented:

“I think I consciously tried not to talk. Then I realized I wasn't talking at all.”

4.5 Discussion

This exploratory investigation with the Conversation Clock demonstrates the influence a social mirror
can have on the group dynamics in collocated conversation. By providing visual cues generated from
the ephemeral audio conversation, we have created a visualization that allows participants to evaluate their own interaction and the history of conversation during that conversation. In this section we will summarize and discuss our findings.

- **People perceive a notable difference between short and long histories.** According to self reporting, longer histories are more effective at altering one's own interaction. However, no noticeable change in gaze, gestures made, time spent leading, turns taken, or length of turns ever corroborated this observation. We speculate if this perceived change does exist, it was something which we did not measure or the difference in what we did measure may be more subtle. Additionally, due to fallibility of self reporting, there may be no real difference in interaction (Bernard et al., 1984).

- **The Conversation Clock encourages participants to remove their gaze from the conversation.** People looked away from the conversation more often. Glances to the Conversation Clock did not replace other glances that occur naturally during a conversation, and participants did not report a loss of quality interaction. As supported by the qualitative data, the social mirror led individuals to become curious of their representation. Some glances to the table are a result of inspecting one's own interaction.

- **Talkative speakers alter the length of their turns while quiet speakers alter the number of turns.** These observations illustrate how the common social mirror can be individually interpreted. The different interpretations influenced their verbal adjustments and interview responses. Quiet participants commented on domination of the entire table, noting it was dominantly one color. Whereas talkative participants mentioned completing rings as a milestone. This difference in observation and perception of interaction demonstrate the principle of small multiples (Tufte, 2001). In the larger picture made visible by history, one can see how they have contributed over the course of conversation. In the outer ring, the Conversation Clock focuses on detail for the recent past, distinctly affecting the current speaker.

- **Participants showed particular interest in their own interactions.** Qualitative feedback indicated that people were most aware of their own interaction and monitored their own contributions. They felt they were speaking too much or too little based upon what they saw. Additionally, some participants evaluated his own contribution and usage of filler words like “umm.”

- **Ordering demonstrates a notable break in period for the Conversation Clock.** Changes in gaze and gesture patterns become apparent when viewed chronologically. Nearly all our measurements indicated significant change. This strong change indicates some break-in period that should be accounted for in all results. One participant's comment almost 30 minutes into the second session underscores this point when he realized, “This IS useful.”
4.6 Continuations

As an exploratory piece the Conversation Clock provided a number of different opportunities and ideas to explore in future work. The Conversation Clock led me to questions such as:

- The social mirror communicates a sense of participation – how might participants demonstrate the value of contribution?
- What about the visualization encourages balance?
- Outside of the actual conversation, the visualization loses context, can the visualization capture and meaningfully annotate a visualization over time?
- How might the visualization change remote interaction and telephone calls, when there is no other visual cues drawing attention?
- What are the long term effects of using a similar social mirror repeatedly?

These questions drove the continuation of my research and inspired the Conversation Votes and Conversation Clusters visualizations and lead to the application of social mirrors to classrooms and branching into therapy. The next chapters document those studies and the plans for future social mirroring applications.
CHAPTER 5

Conversations Votes:
Capturing Anonymous Perception

Having shown the Conversation Clock can communicate passive feedback into conversation, Conversation Votes explored the visualization as a means to provide direct feedback into conversation. The Conversation Votes visualization provides a new channel to empower people who might otherwise remain quiet due to evaluation anxiety, power hierarchies, or the necessities of a large conversation. Extending beyond the Conversation Clock’s depiction of “just the facts,” Conversation Votes becomes an anonymous sounding board to demonstrate the approval of the audience.

While the Conversation Clock affected how people contributed to conversation, it made no distinction between eloquent monologues and random nonsensical babble. The Conversation Votes design allowed participants to make those distinctions as the conversation progresses. Each person indicates agreement and disagreement by pressing inconspicuous buttons. The resulting anonymous feedback is immediately visible, within the small group.

Anonymous feedback proved to be an enabling channel for participants. The study for this interface allowed participants to “debate” contentious topics between group members and every group member was encouraged to express their opinion. People who most used the feedback channel felt the opinions discussed were not adequately representative of their own opinions. With feedback incorporated into the visualization’s display, these same people asserted themselves by speaking more in the conversation. They spoke more, and felt better about the debate.

We designed the anonymous back-channel as a medium for those with less social standing to offer their opinions (Sproull and Kiesler, 1991). In large groups, conference calls, or public meetings there is not always the opportunity for all to speak; participation favors those with higher social rank. The Conversation Votes visualization annotates aural activity with explicit voting feedback on the tabletop display. The resulting public image conveys the group’s opinion of the conversation and denotes salient moments. Additionally, the anonymous votes further lowers the social cost of entry into conversation allowing a wider range of ideas in discussion (Kiesler et al., 1988).
5.1 Design Choices

Conversation Votes builds upon the design choices of Conversation Clock. It incorporates the centrally located visualization, the minimization of movement, and an increased highlight of focus on the most recent time. The design of Conversation Votes happened in parallel with the Conversation Clock and shares the background design process. The truly distinctive aspects of Conversation Votes design were in determining the best way to demonstrate a vote. How much should it change the visualization, how often should a person be able to vote?

In other tools, anonymous feedback has been an effective way to gauge a groups unbiased reaction in asynchronous interactions like elections to synchronous interaction on group support tools (Gavish and Gerdes, 1997; DeSanctis and Gallupe, 1987). It is an effective way to raise minority opinions as the anonymity encourages a broader participation. Conversation Votes focuses on providing that feedback in a face-to-face situation, to augment group understanding when there is a lack of anonymous channels. However, its focus on feedback presents additional visual aspects to incorporate into both the visualization and face-to-face interaction.

Face-to-face interaction uses back-channels to provide richer communication as discussed earlier in Chapter 2: gestures, eye gaze, verbal utterances. None of these are anonymous and directly affect the perception of one’s face (Goffman, 1963). Our design decisions reflect the knowledge that these channels are useful when applied in other areas, but are not applied in the face-to-face environment.

Visual prominence plays an important role in Conversation Votes. As a person makes a point in conversation, the group’s perception changes that person’s representation in the visualization. Points of agreement become more prominent while disagreement becomes less prominent. The visualization mapped prominence to the size of the bars rather than amplitude. Size is one of the most visually distinguishable aspects of visualizations (Bertin, 1983): so the Conversation Votes normalizes the amplitude of all bars to make voting stand out.

5.2 Conversation Votes

Conversation Votes introduces explicit feedback into the social mirror. As a social mirror it provides a real-time common visualization for a group of four participants in a conversation. The following sections describe the visual detail and physical interface.
The Conversation Votes table allows participants to view their conversation with an additional perspective. All members at the table are able to indicate their agreement or disagreement.

**Figure 5.1**: The Conversation Votes table allows participants to view their conversation with an additional perspective. All members at the table are able to indicate their agreement or disagreement.

### The Timeline and History

Conversation Votes presents a structured timeline to highlights the most recent past while summarizing interaction history. The passage of time appears as a sequence of rectangular bars. The length of each rectangle represents the average audio sample for a single second. As shown in Figure 5.1, rectangular samples progress through the table’s center. This center progression shows one minute of elapsed conversation. Its central location on the table provides a detailed view of the interaction most recent in history.

Either side of the center progression depicts older minutes in the conversation (Figure 5.2). Details are smaller and less visible, allowing participants to get a higher level view of who talks, how long people talk, and who received votes. The history stacks accumulate to 16 minutes, each stack showing one minute of samples. The leftmost stack of rectangles in our images indicate the most current minute. Conversation Votes replicates the full history above and below the central progression, making it easily visible from any seat at the table. As each minute ends, all completed minutes slide over to make room for a new minute (Figure 5.3). The timeline and history provide the structure to view the individual contributions that make up conversational cues.

### Contribution and Voting

Conversation Votes shows interaction by noting who spoke when and for how long. Individual microphones monitor each speaker, and unique colors identify each microphone in the visualization.
Figure 5.2: The central timeline dominates the Conversation Votes visualization. Each second, it samples and appends the current state to the leftmost end of the progression. All bars in the progression slide to the right to make room. Dotted rectangles mark the time a vote took place and multicolored bars show any simultaneous speech.

Figure 5.3: On either side of the central timeline a longer history builds up. Each collection of samples represents one minute of time, the most recent picture furthest to the left. Over the course of a minute, the left most collection starts with no samples and grows. It then shifts to the right when completed.
With no votes, the visualization provides a simple color coded view of who spoke in the conversation. Unlike in the Conversation Clock, the amplitude of speech is not presented in the visualization. Initially each sample is of uniform length, favoring no individual.

While the visualization automatically captures a conversation’s aural features, it relies on the participants’ votes to highlight a salient moment and provide feedback. Participants cast positive votes and negative votes (Figure 5.4) at any time during the conversation to indicate approval or disapproval respectively. A positive vote increases the size of the sampled bars (Figure 5.5) while a negative vote has the opposite effect on the same set of samples. Additionally, a positive vote brightens the colored interior of the bar while a negative vote causes the color to fade into the background. The vote influences adjacent bars for visual impact while acknowledging the difficulty in pressing a button at the exact moment an utterance occurs. While a vote will always occur after the exact instance that inspired a participant to vote, a vote affects adjacent bars in both directions under the assumption that 1) the button press follows the most salient moments and 2) salient moments are not instantaneous and continue after the button press.

**Figure 5.4:** The voting button is a handheld plate with a colored circle indicating where to push. The button allowed for discreet and hidden voting. The button provided no tactile feedback, a light press was sufficient to vote and alter the visualization.

**Figure 5.5:** The effect of a positive vote lengthens the otherwise uniform bars.

A person can hold the voting buttons discreetly and press them with little effort. By using his or her two buttons, each listener alters the representation of the current speakers. A viewer sees which speakers provided a greater positive contribution to conversation by examining the full history. Larger and more saturated bars distinguish positive contribution at a glance. However in our pilot study, participants found negative votes hurtful. Some voiced their concern about the animosity created. No
one wanted to end his or her utterance on a negative note when in such a small group. These same participants agreed that it might be much more useful in larger, room-filling, groups and crowds. For this reason, we removed the negative voting button for our full study of small groups. Though we recognize that their utility might be more appropriate for other scenarios.

**Simultaneous Speakers**

Moments of simultaneous vocalization, indicating excitement, agreement, or contention, proved to be among the most salient aural features portrayed by the Conversation Clock (Chapter 4). We incorporated this visual cue into Conversation Votes by showing the two loudest active speakers in each sample bar — though still without the amplitude. The speaker with greater volume determines the color of the outer bar while embedded rectangle represents the second speaker. Previously, overlapped bars would indicate relative amplitude of each speaker. As the bar length is no longer indicative of amplitude, the decision to only show two speakers makes the visualization more legible.

5.3 **User Study**

We set out to evaluate the following questions:

- **Q1** Do less talkative members provide more anonymous feedback?
- **Q2** Does receiving votes change a person’s interactions?
- **Q3** Will positive feedback disrupt the trend to balance conversation?
- **Q4** Does voting convey participants’ views in the conversation?

**Methods**

We gathered 24 volunteers (13 male / 11 female) to meet in 6 groups. Participants consisted predominantly of undergrad and graduate students in engineering disciplines. We asked each group to meet once in the HCI Lab and take part in three debates.

The four participants of a group sat at our rectangular table, two per side on the long side. Prior to beginning the study, each participant clipped a lapel microphone to their own collar before we began calibrating the microphone sensitivity. Participants were also given a single button and told to indicate
their approval and encouragement of the current speaker by pressing it. As there is little sensory feedback when pressing the buttons, all participants first tested a button press before beginning the session.

A full session lasted about 1.5 hours and consisted of three 15 minute mini-session debates. In order to provide a base level of activity for comparison, we conducted the first mini sessions without the visualization projected. In this session, participants were not aware of the visualization's appearance and voted knowing only they were marking positive moments in conversation. The second mini-session began with a demonstration of the previous topic's visualization, an explanation of how conversation depiction, and a live demonstration. We projected the visualization for the duration of this topic. The final mini-session appeared without the visualization to offer a comparison with the baseline.

Topics for each mini-session came from a collection of debate topics for youth debate groups that would be familiar and easily understood by most participants: the minimum age to allow voting, the establishment of national ID cards, and banning smoking in public places. The debate nature of the discussion was to provoke a confrontational style of conversation. Each group received a unique ordering of debate questions that we assigned prior to each session. Participants were free to argue either side of the issue, explore a topic, and switch sides during discussion.

For each mini-session, we logged the aural activity and button presses used to generate the Conversation Votes visualization. Specifically, we logged how often they Lead conversation, how many Turns and the Turn Length, and the number of Votes. These measurements are all straight forward and simply tracked by the Conversation Votes system during conversation. We also generated a measure of Voting Effect to represent the visual "bump" a participant receives. We calculate voting effect as the increase of a participant's graphical rendering from a baseline of no votes. All of these measures were normalized to be per minute for comparison across groups.

To measure individual perspective on the conversation, mini-sessions concluded with the same brief questionnaire. We presented participants with three, seven point, Likert Scales to measure how adequately everyone's viewpoints were Represented, how Comfortable participants were in the discussion, and how much their opinion was Altered due the discussion. Additionally, we asked all participants to notate the degree of contribution each group member made to the conversation.

The second session questionnaire included an additional set of seven point Likert scales to investigate the visualization awareness, visualization accuracy, level of voting anonymity, and degree of altered participation due to the visualization.
As part of the survey following each mini-session, participants we asked participants to estimate the total contribution during conversation. We compared this attribution to the logged lead data to calculate the Estimation Error in all conditions.

Results

Overall, our investigation demonstrates that anonymous voting creates an effective back-channel to enable some, though not all, to better assert themselves in conversation. We expected to aid the quiet individuals to express themselves, however this was not the case. Those enabled by the back-channel were a subgroup that felt the unaugmented conversation was less inclusive of all viewpoints.

To investigate our earlier questions, we classified and divided participants for comparative analysis. Similar to in prior work, we divided groups based on aural participation in the initial session; Heavy contributors spoke more than the leads per minute median value while light contributors spoke below that same threshold. We also included additional divisions on voting and receiving votes. To explore voting, participants we grouped participants into active voters and less active voters to examine how they voted, and we grouped participants into heavily supported and lightly supported based on the voting effect of received votes.

For the statistical analysis, we fit our data to a Linear Mixed Effects Model using Hierarchical Linear Modeling with a repeated condition. A generalized linear model, it is commonly used to address hierarchical data models in social and behavioral sciences when analyzing groups with set hierarchies. Modeling our data as individuals who are a part of a group, we acknowledge individuals are not independent observations and account for the variance that naturally occurs between groups. To investigate Q1–Q4 posed earlier, the model also included variables indicating splits defined above.

As in Table 5.1 Conversation Votes altered turn length across all individuals \( F_{2,241} = 3.68, p < 0.04 \). Follow-up investigation reveals that turn lengths decreased in the final mini-session. Table 5.1 also shows error estimation approached significance, indicating that people are modestly better at estimating contribution when they have a visual representation available. Surveys presented during the visualization session indicated participants were aware of stronger changes than these initial results reveal. They noted that both their own and others’ participation were altered from the previous session (Table 5.2). This does turn out to be the case, but it is only after analyzing changes throughout the sessions when split that it becomes apparent.
We found only minor differences outside of the defining characteristic when making comparisons between the splits overall. Heavy participants tended to speak about 9 seconds more per minute than the less active participants \( F(1,21.1) = 38.83, p < 0.001 \) and take an additional turn every two minutes \( F(1,21.7) = 19.83, p < 0.001 \). Active voters pressed their buttons an additional time every 2 minutes \( F(1,23.0) = 13.47, p < 0.001 \), and heavily supported participants received 10% more increase in visual prominence than did the remaining participants \( F(1,28.2) = 11.69, p < 0.002 \). However, we see more interesting differences when we examine the splits over the course of the three sessions (Table 5.3).

The first split, between heavy and light participants, provides a point of comparison with previous work showing that visual feedback of group activity tends to balance contribution (Section 2.4). Our results with the Conversation Votes supports that finding \( F(2,21.1) = 5.42, p < 0.012 \). Participants noted a change in their debate, stating “[it] more evenly dispersed conversation. I was less likely to interrupt.” Overall, the visualization encourages a more equitable distribution of contribution and balances participation in conversation.
Examining differences between active and less active voters reveals a link to participation levels (Table 5.3). Unlike the balancing seen in the previous split, when active voters can see the results of their votes they drastically increase their lead in the conversation, diverging from the non-voters ($F_{2,21.3} = 5.35$, $p < 0.013$). Looking at Figure 5.6a we see that both groups are essentially equals in leading conversation when no visualization is present, but active voters speak about 30% more than less active voters with the visualization.

These same active voters reported better representation of opinions when interacting with the visualization present ($F_{2,22.6} = 4.45$, $p < 0.021$). Seen in Figure 5.6b active participants were less satisfied than their less active counterparts that conversation was providing a full representation of viewpoints.

Table 5.3: Splitting the participant groups on three different variables — Talkativeness, Voting Patterns, and Voting Support — the linear mixed effects model highlights the differing interaction emerging from our split categorizations over time and across the visualization conditions. Extended versions of this table appear in Appendix B.
Feedback also indicated the back-channel could have been better utilized with a larger vocabulary of signals. One participant stated a negative vote would be particularly useful because he looked when he wanted to move on and “felt someone was talking too much.” Though the channel provides limited feedback, the visualization allowed this group of active voters not only to speak more, but to feel more satisfied with the group discussion.

Though voting enables activity and makes participants feel better about the debate, receiving a vote made little difference in our quantitative results. However, it inspired the most conversation amongst the participants.

“You could see when the others agreed with you, so it encouraged you to continue talking.”

“I could get a visual grasp of argument/conversation successes (i.e. winning others over).”

“[I would] check if others were agreeing with the point presented (not necessarily by me).”

In spite of our participants receptive comments, the last split examining the heavily and lightly supported shows no significant differences, only two notable differences. These notable differences in leads $ (F_{2,21.3} = 2.85 , p < 0.080) $ and votes $ (F_{2,19.1} = 2.79 , p < 0.087) $ fall in line with the above quotes. However, we need a larger testing pool to confirm those receiving votes become more talkative and more apt to vote during the Conversation Votes sessions.

5.4 Discussion

The results demonstrate that a conversation visualization with a voting back-channel can influence conversation and perception. Below is a brief summary of highlights from the qualitative and quantitative results, noting how it relates to our original four questions.

Opening a back-channel: With Q1, we sought to show an anonymous back-channel offered an outlet for those reluctant to speak up. We had expected the light contributors would utilize the back-channel while heavy contributors favored speaking. However, our results do not show a significant difference in the voting patterns of heavy and light contributors.

Instead, active and unsatisfied voters increased their participation over the less active voters with the visual cues (Figure 5.6a) to assert their unheard opinions with the group (Figure 5.6b). Qualitative
feedback also indicates the back-channel helped the group to better craft their arguments by understanding the group's overall sentiment. The back-channel did not serve the purpose we had targeted, but it did create a new medium to better shape conversation contribution.

*Participants strive for balance:* The social mirror encourages people to have a balanced conversation as posed in Q3. Previous work found significant change towards balance when participating in a visualized conversation. Presenting perceived contribution, as opposed to raw data, did not change this effect.

However, we cannot claim that a balanced conversation is a necessary goal of good conversation. A balanced conversation might be a result of silencing a more informed or provocative speaker. Further study is necessary to investigate the definition of quality in conversation, though our results for Q1 and opening a back-channel indicate that the balance did not lessen the quality of group conversation.

*Awareness of Self and Others:* With the visualization, participants reported being more aware of others' contributions. They reported checking the visualization for agreements and approval of points. The visualization became a testing ground for ideas and feedback into one's success within the rest of the group. Participants also reported checking for reactions in response to other people's points. For Q4, we argue the visualization does adequately convey how participation view conversation.

*The Voice of the Voter:* The heaviest voters were less satisfied with the overall representation of ideas. Their voting could indicate pressure toward other topics. As one participant stated, the vocabulary of feedback should cover a wider range of comments. While we removed negative votes for our study, our quantitative and qualitative results support an expanded set of cues. We have shown the visualization provides the necessary back-channel to send cues anonymously, however, the voters desired a back-channel with more than just the positive vote.

*Social Mirror Karma:* In examining Q2, we cannot definitively say receiving votes changed a person's interaction; our numbers are not strong enough to be certain. However, combined with the qualitative feedback, we hypothesize what further work might show.

Heavily supported individuals are more talkative and more active in voting when visual feedback is present. Participating in the conversation and being active in the social mirror seem to correlate with receiving more votes from the remaining participants, in a sense one must give in order to receive. While a participant mentioned the possibility of gaming the system by anonymously voting for himself, an examination of the logs does not indicate it occurred in our study.

From our own observation, we posit that the engagement in conversation influences both receiving and casting votes. A participant in conversation is likely to speak more and vote more when a topic is
close to his or her knowledge or interest, encouraging them to influence the social mirror to support their ideas. Future work should consider measuring prior knowledge and taking it into account.
CHAPTER 6

Application:
Feedback in the Classroom

The social mirror offers insight into a group’s dynamics and the perception of others. A single group mirror provided a channel in which all people are equally represented. With the Conversation Clock persistency of interaction allowed people to judge their social appearance. With the Conversation Votes explicit feedback from the audience provided a more nuanced view in discussion. Based on the Conversation Votes’ feedback, we saw that individuals unsatisfied with discussion would speak up when a back-channel was present. The Fragmented Social Mirror continues this line of research by applying a similar back-channel to a classroom environment.

In this study, we fragmented the mirror into multiple pieces to deal with the different but well-defined roles in a classroom context. In a fragmented social mirror, individuals are not equally represented in the mirror. People can direct their feedback to a specific speaker, individuals have a personal screen, some information can even remain hidden. This was the solution to deal with the unequal time speaking and the need for specific types of feedback such as question asking and question answering.

The Fragmented Social Mirror presented in this chapter was designed specifically for a classroom or lecture scenario. The feedback serves different purposes for the two audiences (lecturer and the masses). For the lecturer, the feedback is a window into the students understanding. It provides a mechanism to gauge the students and alter the presentation on the fly. Additionally, it offers a mechanism to review the lecture after the fact to see where students had the most questions. For the greater audience, people can provide directed feedback and participate in conversation without interrupting the flow. Students can also use it to understand that they are not the only one with questions.

Application Motivation

Students learn more when they actively engage in the classroom [Weaver and Qi, 2005]. However the structure of many classes ensures that the lecturer speaks for at least 80% of the time. Though some
students participate, it’s expected that 5 students out of 40 will dominate any classroom discussion. The lecturer’s awareness of class comprehension is skewed by the students’ many social pressures and the few speaking opportunities. Students try to present a positive image of themselves to their peers. They often avoid volunteering information due to evaluation anxiety, a fear of being judged by others for making a mistake, or being the focus of attention \(\text{(Weaver and Qi, 2005)}\). It’s easy to remain silent. Those students who do speak are generally self-confident or understand the material. However, there is a reluctance to appear too engaged in the classroom. Those who raise the expectations on a group may be ostracized by their peers \(\text{(Parks and Stone, 2010)}\).

This chapter presents an interface prototype designed to encourage student engagement and improve the lecturer’s awareness in the classroom. The prototype, entitled Fragmented Social Mirror, aims to create a new communication channel of anonymous dialog between the instructor and the class. Unlike many previous Audience Response Systems \(\text{[Fitch, 2004; Stowell and Nelson, 2007]}\), Fragmented Social Mirror allows for expressive feedback via text and is accessible at any time during the lecture. In our short pilot observation, students in a large class began to initiate interaction with the lecturer, whereas previously they had only mumbled answers in response to posed questions. In this chapter, I describe the design of the Fragmented Social Mirror in the context of other Audience Response Systems. We also discuss promising initial observations from a classroom pilot study of the Fragmented Social Mirror.

### 6.1 Fragmenting the Social Mirror

The Fragmented Social Mirror provides feedback based on principles described in Chapter 3 however, the classroom setting necessitates a break from the standard social mirror design. A social mirror is a real-time depiction of interaction meant to augment natural face-to-face interaction. It captures ephemeral moments in conversation and brings them into the public view through visualization. In previous chapters, social mirrors displayed abstract visualizations of each individual’s participation in conversation. The resulting display of conversational dominance, non-participation, and turn taking encouraged more balanced conversation. In these social mirrors, one shared visualization of conversation was projected centrally for all participants to see. This form is ill suited for the setting of a large lecture or classroom. There are many more participants involved, and the architecture of the space differs from the spatial layout of small group interaction around one shared table. Furthermore, there is a natural asymmetry in participation due to the lecturer-audience dynamic. This results in less interaction between the lecturer and the audience.
The term “fragmented” in Fragmented Social Mirror refers to the use of individual interfaces for each participant as opposed to one shared visualization and the shortened time component as opposed to the full history present in previous social mirrors. In this setup, each participant accesses a Java applet from their computer or mobile device, while a large public display is presented to everyone. Furthermore, while a traditional social mirror maintains a persistent history of interaction, Fragmented Social Mirror highlights questions and comments that are pressing at the specific moment.

**Design Choices**

The Fragmented Social Mirror design focuses on capturing and reflecting the unheard and unvoiced dialog in the classroom. In addition to the results of the Conversation Votes, we investigated classroom behavior to observe students. We informed our initial design by observing an active and engaged classroom of 100+ students to see what students say when engaged in an active class. The lecturers of these classes were generally rated as among the best in the department. They were engaging during lecture and good at encouraging student participation. To facilitate more participation the lecturers posed a question and waited for responses — thus guaranteeing an answer or a question for clarification. We noted all the types of student responses to better understand what a student wants to say during class. The responses were narrowed down to the following list of categorical responses:

- **Questions**: Students provide new questions based on what has just been taught.
- **Information**: Students add their own connection to outside subjects.
- **Agreement/Disagreement**: Students answering a lecturer’s question.
- **Slow Down/Redo**: Students did not understand the lecturer.
- **Cannot Hear/Repeat**: Students did not hear the lecturer.

After refining numerous sketched prototypes (Figures 6.1 and 6.2), we settled on a simple interface students could use without pulling their attention too far from the lecturer. Thus the input of the Fragmented Social Mirror was used only for capturing one short comment at a time. The history of feedback was only seen on the public display and limited that history to the most recent comments. Additionally, the needs of the lecturer necessitated this type of design. The lecturer needed to be able to read feedback from the hundreds in the audience while still being able to teach effectively. In past chapters, the social mirror was primarily viewed by the listeners (and not the speaker) in conversation because they had more free attention. In this design, the captured feedback of conversation is significantly pared down, so that the lecturer can receive the benefits from the social mirror with
Figure 6.1: Many of the early prototypes for a large group continued to display captured audio in the display. In a large unbalanced conversation such as a lecture, that feedback seemed less important to a participant. We opted to drop this feature in favor of a simpler interface.

Figure 6.2: This prototype allowed participants to send feedback to any speaker and not just the lecturer. However, initial testing deemed this interface too involving. Participants would dedicate much more attention to the interface than to the interaction. However, if there were fewer cues, as in a teleconference, it may still prove beneficial.

minimal attention. Therefore, current comments/questions are displayed so as not to overwhelm the viewers with a long history.

Iconographic Messages

The Fragmented Social Mirror interfaces passes information through icons. These graphics serve to simplify the message so that the lecturer might easily understand the classroom without reading too much content. Based on some informal observation of classroom sessions and some prior work (VanDeGrift et al., 2002), we selected six basic messages to pursue: “I have information/an answer,” “I have a question,” “Yes/agree,” “No/disagree,” “Speak Up,” “Slow Down.” Three researchers independently
drew any graphic that they felt reasonably captured these messages. We combined them into sets for each category, with a total of 5–15 images for each message.

We conducted a survey of Computer Science undergraduates to test our icon designs. A total of 54 Computer Science undergraduates completed our survey. Their feedback identified 17 icons that convey the intended message. Table 6.1 shows all 17. None of the icons for “Slow Down” conveyed an adequate message to the student. We eliminated this message as well as the “Speak Up” messages in favor of a simpler 4-icon interface. Students can use the Information and Question messages with additional text to signal “Slow Down” and “Speak Up.”

**Fragmented Interfaces**

There are two Fragmented Social Mirror interfaces — the student input interface for a computer or handheld device (Figure 6.3) and a larger public screen for the lecturer and audience (Figure 6.4). We provided the public display on a large projection screen at the front of the room, though the lecturer sees the public display on a personal screen. Four different pre-selected icons were designed to categorize student responses in the student interface. The icons represent: Information, Questions, yes/agree, no/disagree. We selected these four categories based on our initial classroom observation of common student interaction. We designed 5–11 sketches for each of the four categories. A survey completed by 54 undergraduates was used to select the most representative icon for each category and to provide feedback for the selection. Of the four categories or signals, the Information and Question signals can be augmented by a 40 character message. The short messages allow students to clarify their questions or possible answers when there is no opportunity to speak while the yes/no buttons allows students to answer simple questions quickly.

Students use the input interface in Figure 6.3 to send their message to the public display shown in Figure 6.4. The Fragmented Social Mirror sorts all messages on the public display by their associated icon to increase legibility for the lecturer. The speaker can look up and see many questions that need to be addressed or they can glance over answers that students provided via the display. The icon with the most messages moves to the top of the screen with a larger icon. The most recent message of this icon appears at the top of that icon in white text set against the black background. As a message ages, it fades to grey before finally disappearing after a predetermined delay. For icons with multiple messages, a count is displayed to the left of the icon.

The public display shows only the most recent messages on the screen. After one minute, a message is cleared. The rationale for the simplicity of the design was two fold: 1) we did not want the lecturers
The icon represents confusion or the desire to ask a question

The icon represents that an audience member has an idea or new piece of information

The icon represents disagreement or a negative sentiment

The icon represents agreement or a positive sentiment.

The icon represents that the lecturer needs to increase their speaking volume.

Table 6.1: 54 individuals rating our icons produced these icons as the most indicative of their intended messages. The results informed our final selection of icons seen in the interface (Figure 6.3). One other message, to indicate the speaker was moving too fast, produced no positively rated icons. All bar graphs appear at the same scale.
Figure 6.3: The input interface provides a simple design for a small window. It allows a student to send one of the four preselected messages and text (when applicable) from their computer to the public display. It does not require a lot of attention in order to encourage more attention on the classroom activity.

Figure 6.4: We projected this larger display at the front of the classroom while the lecturer could see a smaller personal copy. The screen anonymously captures questions and comments that students might not otherwise ask. The public display groups messages by icon and highlights the most recent feedback from the class. To the left of each icon, a counter indicates multiple messages of the same type, particularly useful when using the yes/no feedback buttons. As we saw in our pilot study, it enables the students to initiate a dialog with the lecturer.

tobe confused or overwhelmed by reading old questions from a prior part of the lecture and 2) if a question goes unanswered and disappears, this removal may encourage a student to verbalize the question in class or to repost it. One of our main goals is to encourage more class interaction. If a student can "see" that they are not alone in their confusion, they may be less apprehensive to speak out and ask a question.

Once a student sends a message they are blocked from sending additional messages for a brief period (10 seconds in our pilot) to discourage social chatter and channel monopolization. While there is some room for abuse as with the backchan.nl system, where some users voted up questions for humor (Harry et al., 2009), the public availability of the channel is ultimately at the discretion of the lecturer.
<table>
<thead>
<tr>
<th>Session</th>
<th>Students</th>
<th>Computers</th>
<th>On FSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>108</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>125</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>112</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>102</td>
<td>23</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>73</td>
<td>14</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 6.2: This table shows the number of people in the classroom for each session and the number of visible computers. Though we requested that students bring their computers to the sessions with the Fragmented Social Mirror, the number of computers remained essentially unchanged.

6.2 Pilot Study

We conducted a pilot study to investigate the Fragmented Social Mirror in the classroom. We began by observing the participation levels before the introduction of the Fragmented Social Mirror and again with the Fragmented Social Mirror in place. For this, we observed a required first year course with roughly 180 registered students at the beginning of the semester. The instructor was not affiliated with our research team. We observed a total of six course sessions: three without any augmentation, and three with the Fragmented Social Mirror. During observation, an average of 100 students were in attendance in each session, though there were fewer students in the final sessions. Given the large class, not many students had the opportunity to speak, and most did not. A summary of the attendance is visible in Table 6.2.

Prior to testing the Fragmented Social Mirror in class, we sent a pre-survey and described the use of the Fragmented Social Mirror. The survey inquired about the student’s comfort level while participating in class versus their smaller discussion sections. Feedback from the survey confirmed that students are not comfortable asking questions or asking for clarification during class, though they are more comfortable asking in their smaller recitation sections. Similarly, they recognize that they do not participate during class (Table 6.3).

Our initial observations showed little interaction between audience and lecturer over the course of three 50-minute sessions. The only activity from the audience was in response to questions posed by the lecturer. In reference to a proof “n is divisible by what?” or “What is the cardinality of set Q?” The class averaged about four responses per class. The students initiated zero interactions themselves, five

\footnote{The observations with the Fragmented Social Mirror saw fewer students due to the scheduling in the semester. A midterm, review sessions, and the final day to drop a class occurred between sessions 4 and 5.}
Table 6.3: Students reported they were uncomfortable asking questions in class, but more comfortable asking questions in smaller recitation sections. They are aware that they do not participate in class and they do not ask the questions that they have. There is a reluctance amongst the students to speak up and participate. 23 students responded to our survey; all graphs have the same scale.

of the twelve responses were general indefinite murmurs from the class, and two responses involved raising hands. Various sets of 1–3 unidentified students spoke up to answer the remaining six questions.

We tested the Fragmented Social Mirror in three class sessions and found the students were pro-active in using the system. In the classroom, the lecturer used a central projection screen to work through problems by hand while a smaller screen displayed the public display to the right of the larger screen (Figure 6.5). At the lecture podium, the lecturer also had a copy of the public display available during the class activity.
We tested the Fragmented Social Mirror in a large lecture hall with three projection screens. Typically, the instructor repeated the same material on all three screens. For the study, the Fragmented Social Mirror replaced one screen during class-time. The lecturer also has a private screen showing the public display.

With the system in place, students initiated dialog with the lecturer 11 times, compared to zero without the system. When on topic, students used the Fragmented Social Mirror to ask questions of the lecturer, keep the lecturer from moving on too quickly, and to answer any questions the professor posed. Some example dialogs appear below, Appendix C presents the full dialog. Figure 6.4 summarizes the participation in each of the 6 classes. Most of the on-topic dialogs either began with or contained a question for the instructor. They lead to discussions with the instructor and information to enrich the class. However, there were also many off-topic messages. These messages were irrelevant to the class topic and were used to draw the attention of other students away from the lecture.

<table>
<thead>
<tr>
<th>Session</th>
<th>Questions Posed by</th>
<th>Dialog Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Instructor</td>
<td>Student</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 6.4: The Fragmented Social Mirror encouraged students to initiate questions and dialog with the instructor. Many dialogs began with or contained questions while others were responses to indicate comprehension of the material. The anonymity of the interface also encourages unrelated dialog in class. A dialog in the Fragmented Social Mirror could include multiple messages — thus, # dialogs (# messages) includes both the count of interaction instances and total messages.
Example Dialogs

Excerpts from the Fragmented Social Mirror sessions appear below. In this first example, students requested information that the professor was not trying to teach but established an interesting aside on history related to the lesson:

**Instructor**: [discussing the Karatsuba Algorithm]

**: <What is Karatsuba?>

**Instructor**: Karatsuba is the guy who invented it, Anatolii Karatsuba. [Instructor continues with a bit more history].

As another example, the student's lack of understanding prompts the student to ask for clarification on calculating tree height.

**Instructor**: [Explanation of tree depth]

**: wow

**: What is the height again?

**: The maximum depth of the tree. You can count the Levels by generation

**: Not 5? It's max level and not count? 

**Instructor**: Yeah, it's 4 not 5 …[continues on 0 based counting]

**: In cs you start counting at zero :)

However, with the addition of initiating comments, there was also an increase in comments solely intended to draw attention away from lecture. These messages, often had nothing to do with the lecture. They tended to come in bursts in order to dominate the activity on the public display for a short time. As an example of such a burst:

**: HATE HATE HATE HATE HATE HAT

**: I DONT CARE WHAT THESE CHICKS SA

**: I DONT EVEN LOOK THAT WAY

**: EVERY TIME I WALK IN THE CLU

**: THEY HATIN ON ME CUSE THEY KNO I LK GOO

This type of interaction was most prevalent in the second session. The lecturer was inclined to read them to see if they were relevant to the class. To indicate the message had been seen, the lecturer then
laughed, if it were funny, or stated “I don't know what this means.” However, the increase of messages also meant that the lecturer was more likely to miss relevant exchanges, such as a student was asking for help:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>can you draw the picture for the tree after applying rule 3</td>
</tr>
<tr>
<td>2</td>
<td>:(</td>
</tr>
<tr>
<td>3</td>
<td>+1</td>
</tr>
</tbody>
</table>

After the Session

We had only planned to gather initial observations to refine the system in these first sessions; however, the instructor was excited to see the students participating and invited us to return with the system for further studies. After the lectures, she indicated that it’s always been hard to get this many students to say anything, even with encouragement. The simplicity of the display was also deemed useful, as she could read the questions with a glance. Additionally, the asynchronous nature allowed students to ask their questions while she was still explaining — thus allowing her to work the question into that explanation or come back to it later. Student feedback indicated the device was useful as they “didn't have to try to get the professors attention” by raising a hand from the back of the lecture.

Students also saw the benefit of the interface, and felt it was easier to participate in the classroom (Figure 6.5). However, they recognized the difficulty of maintaining order in the anonymous display and provided suggestions to keep the interface on topic. One such suggestion was to make the display semi-anonymous. That is, publicly anonymous, but retain the identity on the lecturer's display. In this way the instructor could call out any abuse of the display, while protecting the identity of any others who were uncomfortable commenting in front of the class. A similar suggestion would simply log the identities for review after class.

6.3 Discussion

The Fragmented Social Mirror touched the surface of integrated feedback in large discussions by allowing controlled feedback that could express a student’s question or comment. As we mentioned in related work, the field can be divided into the high expressivity interfaces that require more attention
The system encourages participation.

The system encouraged my own participation

The system encouraged me to ask questions

The system encouraged my peers to ask questions

The system made me feel connected to the lecture

The system made me feel connected to the presenter

The system made me feel connected to other audience members

The icons were clear in their meaning

The use of text entry was clear in its meaning

The system was easy to use

The feedback from others helped me to understand the lecture material

The feedback from others made the lecture more enjoyable

The system is a worthwhile addition to the lectures and discussions

**Table 6.5:** Students reported the Fragmented Social Mirror encouraged participation and connected them to the lecture. Students reported the lectures being more enjoyable and saw the feedback as a worthwhile addition to the lectures. All graphs appear to the same scale, 16 students responded.

and the low expressivity interfaces that do not allow students to indicate their questions. Fragmented Social Mirror bridges this continuum.

In our sessions with the Fragmented Social Mirror, the classroom dialog was more involved. The lecturer felt like she was talking to people rather than at people while the students took a more proactive role in directing conversation to points that were not understandable. With 100 students, evaluation anxiety limits the individuals willing to speak. We have shown that anonymous feedback can break the barrier and include more students.

Our interface was not perfect, the classroom sessions revealed that while anonymity opens opportunities for discussion, it must be tempered in some manner. In both the second and third sessions, some individuals engaged in the public conversation by adding potentially disruptive comments that lead to legitimate questions not being discussed. In this prototype, we did not prevent this type of interaction. Perhaps the lecturer should be able to flag such comments during conversation to lock out individuals,
there could be a moderator, or the lecturer should be able to identify individuals after the classroom session and deduct points in some manner. While it is tempting to simply allow the class to moderate itself, it creates an additional task that draws them away from the learning process.

The classroom sessions also made it apparent that the positive and negative responses should be more flexible. Though students did use them as feedback for the lecturer, they also use the agreement check-mark to indicate a “me too” when other students raised a question. Others adopted a convention of adding a “+1” as seen in the example conversations. A revamped system might allow a student to indicate “me too” and ensure the question stays visible long enough for the lecturer to see the question.

Our surveys underscore the need for large classrooms to tap into technological back-channels. Students know that they do not participate in large classroom settings even though they have questions. They are not comfortable asking questions in such a large group. The survey after the use of our system shows that students felt encouraged themselves and the class to ask questions, they found it made the lectures more enjoyable, and it was a worthwhile addition to the lecture.

While our system was successful in engaging students and encouraging participation, we acknowledge it has limitations. Our study was a small study to test our conception of anonymous feedback. The system, with refinements, should be further tested over a longer term and in multiple classrooms to better judge the effectiveness on the learning experience rather than our focus on just the classroom engagement. Many untested facets of the interface can be further explored. We advocate anonymous feedback based on the premise of evaluation anxiety, though we have not yet tested the effects of allowing or enforcing identity in the Fragmented Social Mirror.

6.4 Limitations

The Fragmented Social Mirror indicates the use of text based anonymous feedback has potential for promoting engagement in the classroom. However, this pilot is not a definitive study. The system needs further testing long term and in multiple classrooms. Untested facets of the interface can be further explored. We advocate anonymous feedback based on the premise of evaluation anxiety, though we have not yet tested the effects of allowing or enforcing identity in the Fragmented Social Mirror. We limited the interface to four icons for simplicity, but how many and which are the best to employ? From the limited use we have seen, there is already support for allowing directly up-voting other questions and information in the display.
The Fragmented Social Mirror interface received a positive response from both students and lecturer. Our initial study highlights the use of anonymous signals in large classroom has potential to draw in more active participation of the students and audience.
Distorting Feedback:
Manipulating the Mirror

We have seen shared real-time visualizations of group conversation influence the behavior of individuals in small groups. Previous chapters and numerous researchers have designed visualizations depicting a speaker’s contribution in face-to-face co-located interaction (Bachour et al., 2008; DiMicco et al., 2004; Kim et al., 2008). Repeatedly, studies have repeatedly shown that visualizations of face-to-face participation encourage a balanced conversation. This change in dynamics can be attributed to Goffman’s theories on face (Goffman, 1959). Visual feedback, in the form of visualization, can act as a moderator to encourage balance with other group members, but can that contribution be actively shaped?

This chapter investigates a more nuanced view of the social mirror’s influence on group dynamics. Specifically, I contrast the use of accurate feedback versus distorted feedback with the Conversation Clock by distorting the feedback to indicate a participant spoke much more than in reality. In this chapter, I first demonstrate four distortion strategies and their effectiveness at misleading the viewer before applying them to group conversations. In the presented studies, I show that people will trust and accept a significantly distorted social mirror as a representation of conversation. However, only a small change in participation can be attributed to this distortion. In dyadic conversation, distortion suggesting a person speaks 60% more produces only changes the conversation balance by 8%. In groups of three the same shift was not detectable. I conclude that the driving force of a social mirror is not directly tied to the accuracy of the visualization it displays.

7.1 Distortion Strategies

The distortion in the Conversation Clock aims to make people appear more talkative and dominating. In a way, it artificially increases the salient moments that individuals noted in Chapter 4 by showing more contribution throughout conversation and generally make the visualization dominated by a single speaker’s color. We began with a list of four potential distortion strategies for testing in a pilot
study: amplitude (volume), speed (sample rate), color brightness of rectangles, and color replacement in past samples (i.e. swapping the color of rectangles from other participants’ assigned colors to the emphasized person’s color). Low, Medium, and High conditions of each strategy were tested. Descriptions of each strategy follow while simplified renderings are shown in Figure 7.

- **Speed** changes the sampling rate of the Conversation Clock. When the emphasized person leads the conversation, the sample rate increases and more bars are drawn for this person. Low through High conditions vary the sample rate to a lower and higher frequency, respectively.

- **Amplitude** increases the length of the speaker’s rectangular bars. This implies the speaker was louder throughout the conversation. Low through High Conditions vary the amplitude multiplier with low being a slight increase to high being the largest increase in size.

- **Brightness** emphasizes one participant by decreasing the color brightness of all other participants. Low through High Conditions vary the percent of reduction in brightness with low being a subtle difference to high being the largest difference in brightness.

- **Color Replacement** changes the rendering of the person speaking in previously drawn history. The emphasized person’s color is used to recolor its neighbors 30 seconds after they are drawn. Low through High conditions vary the number of neighboring bars affected: a low condition will affect the two nearest bars on both sides whereas the high condition changes four bars on both sides.

The distortion strategies leverage natural deficiencies in visual cognition such as change blindness and awareness blindness (Simons et al., 2000; Simons and Chabris, 1999). People cannot attend to all of their surroundings concurrently. Therefore, many of the changes remain unnoticed. In this setting, participants attend to the others at the table, their own discourse, and the table visualization. They are less likely to notice the distortion as it happens. In the case of color replacement, they do not notice the distortion as in changes right in front of them. As feedback from an incorrect visualization is apt to be ignored, we designed the graphical distortions to be subtle and trusted in the short term, yet largely suggestive of dominant participation in the long term.

For our later analysis, we label participants “overemphasized” when they are targets of a distortion strategy. Their contribution is over represented in the visualization. In the same condition, their discussion partner is labeled “underemphasized” as their contribution is visually lessened as a portion of conversation.
Figure 7.1: For the same conversation, the above images depict renderings for the blue speaker under (A) Normal, (B) Speed Emphasis, (C) Amplitude Emphasis, (D) Brightness Emphasis, (E) Color Replacement Emphasis.
7.2 Experimental Studies

We designed a set of three studies to investigate the effect of using distortion in augmented face-to-face conversations. The first provides a pilot study of the distortion conditions. Participants viewed a pre-recorded conversation alongside the visualization. From this study we determined which conditions were effective at misrepresenting balance to the viewer while still being trusted as an accurate depiction of conversation. The second study examines these distortion strategies applied to a group setting with active participants. The third study simulated a remote conversation and removed all face-to-face cues from conversation. In this study, we show that social mirrors can be inaccurate and still produce the change toward balance.

Hypotheses

We began with the following hypotheses:

**H1** A speaker will perceive they are contributing more when their contribution is emphasized via distortion. People will defer to the visualization to gauge their participation in conversation rather than trusting their instinct or memory.

**H2** Emphasizing a speaker's contribution decreases participation from that speaker. Prior work indicates talkative people are conscious of appearing to dominate the conversation and specifically cited the visualization's role in their behavioral change.

**H3** When emphasized, people will report feeling pressure to speak less due to the public feedback. In prior work, awareness of the visualization affected conversation. With distortion applied to an individual, the same awareness of dominating conversation should happen sooner and more often. In addition to changing behavior in H2, we expect that the increased dominance will manifest as a social pressure attributed to the visualization.

7.3 Pilot Study

We began by testing the four distortion strategies to narrow the experimental conditions for a larger study. We predicted our intended strategy would lead a viewer to overestimate the amount of time the distorted individual spoke without the viewer being aware of that distortion. For this pilot study, our selection of distortion strategy was chosen based upon participants' written estimations of participation and Likert scale feedback.
Participants were presented with 4.5–5 minutes of video recorded conversations selected from interview segments of “The Daily Show” while the Conversation Clock visualization was projected synchronously on the table. The video was displayed on a screen opposite the participant. Prior to the study, each video was hand-coded for speaker participation to provide the volume and turn-taking parameters for the Conversation Clock. The interviewer or the interviewee from the video clips were randomly selected to be overemphasized by distortion.

Participants observed a total of 13 total conditions: a Low, Medium, and High condition for each of the four distortion strategies and one control condition. The Low condition was a subtle distortion and would only be noticed if one paid close attention. The High condition was extremely distorted and it was easy to notice they were distorted and the Medium condition fell in the middle. The conditions were randomly shuffled, however the High conditions were never adjacent.

At the end of each condition, the study participants estimated the proportion of time each person spoke during the conversation and indicated the accuracy of the Conversation Clock’s depiction of conversation. The variable, overestimation, was measured as the percent error relative to the distorted person’s speaking time. Eight people participated in the pilot study (5 males, 3 females). They were told all visualizations would differ, though the distortion strategies were not revealed until after the study. The session lasted approximately 1.5 hours, participants were remunerated with gift certificates to Amazon.com.

Pilot Study Results

Figure 7.2 shows the results of the participant estimations as reported in the surveys. All four strategies showed a tendency to deviate from the control condition estimates. We examined the conditions that maximized overestimation for each strategy (High Amplitude, High Speed, Medium Brightness, and Medium Color Replacement) and found that they differ from the control condition via pairwise t-tests. All paired t-tests showed significance (p < 0.04): with our small sample of 8 individuals the effect size is large enough to indicate the conditions were effective.

Aggregate Likert scale data indicated that no distortion strategies were misleading. Individuals voiced a few concerns that varied over the conditions. With the amplitude strategy, some participants remarked “the bars seem bigger” while in the brightness condition some participants noted, “It’s harder to see [the underemphasized speaker’s] color.” The High Color Replacement condition was the only condition to

---

1 A popular comedy-news program
elicit any extreme distrust vocally; a single individual commented “I don’t know how, but it’s wrong. It’s just wrong.” He further explained that he had explicitly watched the bars as they appeared to determine what was going on, and he found no error.

The pilot study demonstrated these distorted social mirrors were trusted and accepted as accurate in a real time situation. As participants reported all strategies to be accurate, we chose the distortions that maximized the overestimation. The two distortion techniques with the highest overestimation percentage were Color Replacement and Amplitude. We applied these two techniques in the following group study. To conserve space, in some charts we refer to the emphasized and underemphasized as CR+ and CR- respectively for Color Replacement and AM+ and AM- for Amplitude.

### 7.4 Group Conversation Study

This group study investigates how our two distortion strategies affect a real-time conversation. We solicited groups of friends to leverage their familiarity. We explained the Conversation Clock visualization, and the participants were given time to familiarize themselves with the colors, time structure, and animation. They were not told about the distortion of the representation by emphasizing individuals until after the study concluded.

A full session consisted of eight 10-minute conditions followed by Likert scale questions. Conditions included a no visualization condition (No Vis), a traditional Conversation Clock visualization, three Amplitude distortion visualizations, and three Color Replacement distortion visualizations. We randomly ordered the visual conditions such that the same person was never emphasized twice in a row. The 10-minute conversations were based on hypothetical questions selected from Gregory Stock’s “The Book of Questions,” and were assigned randomly to the experimental conditions. Participants
were informed they could change the question however they liked or migrate onto tangential topics. Two example questions follow:

- If you were able to live to the age of ninety and retain either the body or the mind of a thirty year old for the last sixty years of your life, which would you choose?

- If you could spend one year in perfect happiness but afterward would remember nothing of the experience would you do so? If not, Why not?

We gathered 12 groups of three people (10 male / 26 female) for the sessions. Participants were undergraduate students, graduate students, and staff from the local university. Full sessions lasted between 1.5 hours to 2 hours, and participants were remunerated with gift certificates to Amazon.com.

**Interaction Measures**

For each person, we recorded the duration of speech, the number of turns taken, the length of those turns, and the degree to which an individual was emphasized. This data was automatically captured and normalized by the Conversation Clock. We also collected the following qualitative data with a Likert scale survey:

- Q1 I found that I spoke ___ compared to others. (Less / More)
- Q2 I found that I spoke ___ than usual during conversation. (Less / More)
- Q3 I found it ___ to make my viewpoint known in conversation. (Easy / Difficult)
- Q4 I felt pressured by others to speak more or less. (Agree / Disagree)
- Q5 I felt pressured by the visualization to speak more or less. (Agree / Disagree)
- Q6 Others were affected by the visualization. (Agree / Disagree)
- Q7 I felt others understood my viewpoint. (Agree / Disagree)
- Q8 I understood other participants viewpoints. (Agree / Disagree)
- Q9 The conversation was natural. (Agree / Disagree)

**Group Study Results**

Overall, people measured their participation based on the Conversation Clock visualization. However, we did not detect additional behavioral change that could be attributed to the distortion. For analysis, we applied a repeated visualization conditions accounting for the data collected as groups rather than independent observations.
As in prior work, we categorized our participants as Talkative and Quiet. We labeled individuals based on their contributions during a preliminary No Vis. The first column of Table 1 demonstrates how our two participant divisions differed from each other. The column highlights the intuitive differences between the Talkative and Quiet participants: they differed in how much they lead conversation ($F_{1,23} = 14.12, p < 0.001$) and how long they speak in a turn ($F_{1,23} = 11.78, p < 0.002$). Talkative individuals speak more overall. Specifically, they speak more per turn than their Quiet counterparts. While the designation of Talkative and Quiet was made at the onset of the study with no visualization present, we found the labeling held through all sessions - not just the No Vis trial.

The remaining columns of Table 1 show where the visual conditions made an impact. The second column compares the changes made across all the distortion conditions. The feedback provided with Q1 (I found I spoke [less/more] than others) changed significantly across conditions ($F_{5,242} = 5.39, p < 0.0001$). The third column compares how the Talkative and Quiet individuals changed differently across the conditions. Significance in Leads ($F_{2,48} = 2.7, p < 0.02$) and Q1 ($F_{5,242} = 2.33, p < 0.04$) indicate that the Talkative and Quiet participants’ reactions to the distortion conditions differed (Figures 7.3 and 7.4).

Though Talkative and Quiet participants alter their speech patterns differently throughout the conditions, their changes do not clearly indicate a connection to the distortion conditions. Examining Figure 7.4, the tendency to balance conversation is still apparent. Overall, the Talkative speak less with a visualization present, and the quiet speak up. We also see that different individual conditions are more effective for Talkative and Quiet participants. That is, Talkative participants were most affected
Figure 7.3: Talkative and Quiet people react differently to the presence of the visualization. However, throughout the conditions only AM- differed from the control condition for both Talkative and Quiet people.

Figure 7.4: People perceived that they were talking less and more than usual in parallel to their emphasis conditions. The Quiet participants were more in tune with the emphasis, whereas the Talkative generally perceived themselves as more subdued in the presence of the visualization.

by the Color Replacement condition. Oddly, we also saw Talkative members spoke less when the other participants were emphasized via amplitude. Conversely, Quiet participants spoke up when others were emphasized (Amplitude and Color Replacement), though no effects were detected in their own emphasized conditions.

Analysis of Q1 suggests the perception of the visualization was altered by the distortion. Figure 7.4 shows that the perception of Quiet participants is more directly influenced by the distortion conditions than the perception of Talkative participants. Quiet individuals report they speak more when emphasized.
and less when underemphasized. The Talkative participants’ highest perceived contribution is in the No Vis trial. The presence of the visualization appears to lower the amount they think they speak.

**Group Study Discussion**

Distorted visualizations impacted the perception of behavior. In both the pilot study and this study participants accepted the visualization feedback as a faithful depiction of conversation. Combining results from this study and the pilot, we were able to validate H1: a speaker will perceive they are contributing more when their contribution is emphasized via distortion. In the pilot study, each of the distortion strategies skewed the viewer’s estimation of speech contribution. In the group study, self report estimates coincided with their respective visualization condition. Participants, particularly the quiet, felt they spoke more when emphasized and less when underemphasized.

Though distortion effectively altered perception, very little measurable effect was detected in conversational patterns. H2 stated that emphasizing a speaker’s contribution results decreases participation from that speaker. We predicted the distortion would mitigate participation and allow other participants to speak more. With the amplitude distortion, some participants were aware of being distorted. Most commonly, people thought it was a miscalibrated microphone. Some participants adjusted the microphone or ask the experimenter to check the audio controls. This issue hadn’t arisen in the pilot where the visualization reflected the contribution of the television speakers. In this group study, participants were more concerned about the details of their own appearance.

H3 stated that emphasized people will report feeling pressure to speak less due to the public feedback. We could not prove or disprove this hypothesis. It is not clear whether the distortions as rendered in the visualization provided any additional pressure to change one’s behavior.

**Post-Study Concerns**

During the experiment, it became apparent that selected groups began to pay less attention to the visualization as the experiment progressed. Post-experiment analysis identified three concerns with the study setup after the fact: group familiarity, topic interest, and group size.

As the recruited groups had a history of conversation with their group members, participants could better judge the degree of changes made to conversational patterns. However, the potential benefit was lost as familiar group dynamics took over. Combined with the interesting topics for discussion,
the groups were comfortably "shooting the shit" as one participant said. The visualization became just a decoration to some, who noted "I didn't really look at it that much." If the conditions had not been done in one sitting, we may have seen more distinct changes in participation.

The group size and the short study duration influenced the study results. As each session was limited to 10 minutes, a fully equal or balanced conversation would allow all members to speak no more than 3.5 minutes. In that time, participants may not have the time to notice and adjust their behavior before moving on to the next condition.

To further explore distortion techniques, we designed a follow up experiment to focus on the visualization and the effect of that distortion. We simplified the study design and focused on the visualization without other collocated cues such as eye contact, gestures, and facial expressions. To accomplish this, we adapted the Conversation Clock so that it could be used between participants that are in remote locations. The participants did not know each other prior to the study. As a result, participants could focus on the visualization for cues rather than just past interaction history.

7.5 Remote Conversation Study

The remote study of the distorted Conversation Clock differed from the group study in the following ways:

1. Participants were remote and could not see each other.
2. Groups were reduced from three to two people.
3. Condition were lengthened to 15 minutes.
4. Speed was used as the distortion technique.

During a session, participants sat in different rooms with a monitor in front of them. The Conversation Clock visualization was rendered on the monitor with each participant seeing the same visualization.

Following the introduction to the Conversation Clock, a full session in this study consisted of fifteen minute conditions: No Vis, Normal, Emphasized, and Underemphasized. For this study we distorted conversation with Speed distortion to forgo the "calibration" concerns of Amplitude distortion while retaining a distortion strategy that changed the visualization in the moment. Since Speed was both effective and never received any negative feedback we chose to apply it to this final study. Conditions were randomly ordered for this session as in previous sessions.
The 15 minute conditions used the same questions from the previous study though each conversation had three questions to ensure there was always something to talk about. Participants were free to answer the questions in any order or deviate from questions as a topic progressed.

Thirteen groups of two participated in this study (10 male / 16 female). Participants were drawn from a similar population at local university, though all were new participants. Full sessions generally lasted about 1.5 hours. Participants received gift certificates from Amazon.com for their participation.

**Interaction Measures**

Similar to the group experiments, we recorded the aural participation in terms of amount of speech, number of turns, length of turns, etc. The survey at the end of the study was slightly modified to fit the new two person design from the original group study. The questions following each condition appears below:

- **Q1'** I spoke ___ than others (Less / More).
- **Q2'** I spoke ___ than usual (Less / More).
- **Q3'** My partner pressured me to speak more or less (Agree / Disagree).
- **Q4'** The clock pressured me to speak more or less (Agree / Disagree).
- **Q5'** My partner was affected by the clock (Agree / Disagree).
- **Q6'** My partner understood my viewpoint (Agree / Disagree).
- **Q7'** I understood my partner’s viewpoint (Agree / Disagree).
- **Q8'** The conversation was natural (Agree / Disagree).
- **Q9'** I was aware of my depiction in the clock (Agree / Disagree).

**Remote Study Results**

This study demonstrates a change in participation as a function of the distortion condition. People adjusted the total amount of their speech when distorted with speed; however, the change was small compared to the degree of distortion in the conditions. Analysis indicates most of this change occurs early and stabilizes afterwards.

The second column of Table 7.2 indicates the emphasis condition significantly affected overall contribution to conversation ($F_{2,48} = 4.2, p < 0.02$). Figure 7.5 makes this effect apparent: emphasized people speak less and underemphasized people speak more. However, the difference is minimal. The
For the emphasized condition, these minutes are consistently lower than the rest of the minutes in that, the progress in an essentially parallel manner. Figure 7.7 demonstrates deviations in minutes 6-8. For the emphasized condition, these minutes are consistently lower than the rest of the minutes in

Table 7.2: The above table shows the results of a repeated measures analysis comparing between conditions, the Talkative/Quiet split, and the two combined. Extended versions of this table appear in Appendix E.

![Bar chart showing Emphasized, Normal, and Unemphasized conditions over time with emphasis on the change between emphasized and underemphasized conditions](image)

Figure 7.5: The remote study demonstrated based on emphasis condition played a significant role in contribution to conversation ($F_{2,48} = 4.2, p < 0.02$). The change between emphasized and underemphasized aggregate results across all participants show an emphasized individual will speak 1.9 seconds less in a minute than when in the corresponding underemphasized condition ($t_{25} = 2.8, p < 0.01$). This small difference was the result of a 60% distortion of the chosen speaker. For every 10 seconds of time an emphasized individual spoke, the speed distortion would render their appearance as 16 seconds of speaking.

Exploring the conversational balance on a minute by minute basis (Figure 7.6), the conditions differ over time ($F_{2,113} = 8.166, p < 0.003$). Specifically, the relative balance between emphasized and underemphasized shifts and becomes stable at approximately the sixth minute of conversation. Beyond that, the progress in an essentially parallel manner. Figure 7.7 demonstrates deviations in minutes 6-8.

For the emphasized condition, these minutes are consistently lower than the rest of the minutes in

<table>
<thead>
<tr>
<th>Measure</th>
<th>Talkative/Quiet</th>
<th>Condition</th>
<th>Cond &amp; Talk/Quiet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F_{1,24}$ Sig</td>
<td>$F_{2,48}$ Sig</td>
<td>$F_{2,48}$ Sig</td>
</tr>
<tr>
<td>Leads</td>
<td>9.8 $p &lt; 0.005$</td>
<td>4.2 $p &lt; 0.02$</td>
<td>0.5 $p &lt; 0.6$</td>
</tr>
<tr>
<td>Turns</td>
<td>0.6 $p &lt; 0.5$</td>
<td>16.2 $p &lt; 0.0001$</td>
<td>0.7 $p &lt; 0.6$</td>
</tr>
<tr>
<td>Length</td>
<td>6.4 $p &lt; 0.02$</td>
<td>1.3 $p &lt; 0.3$</td>
<td>0.5 $p &lt; 0.6$</td>
</tr>
<tr>
<td>Q1'</td>
<td>6.2 $p &lt; 0.02$</td>
<td>0.8 $p &lt; 0.5$</td>
<td>0.5 $p &lt; 0.6$</td>
</tr>
<tr>
<td>Q2'</td>
<td>1.2 $p &lt; 0.3$</td>
<td>1.6 $p &lt; 0.3$</td>
<td>0.2 $p &lt; 0.8$</td>
</tr>
<tr>
<td>Q3'</td>
<td>0.2 $p &lt; 0.7$</td>
<td>0.5 $p &lt; 0.7$</td>
<td>0.2 $p &lt; 0.9$</td>
</tr>
<tr>
<td>Q4'</td>
<td>1.0 $p &lt; 0.4$</td>
<td>0.4 $p &lt; 0.7$</td>
<td>0.4 $p &lt; 0.7$</td>
</tr>
<tr>
<td>Q5'</td>
<td>2.0 $p &lt; 0.3$</td>
<td>0.6 $p &lt; 0.6$</td>
<td>0.8 $p &lt; 0.5$</td>
</tr>
<tr>
<td>Q6'</td>
<td>0.5 $p &lt; 0.5$</td>
<td>0.4 $p &lt; 0.7$</td>
<td>0.4 $p &lt; 0.7$</td>
</tr>
<tr>
<td>Q7'</td>
<td>0.1 $p &lt; 0.7$</td>
<td>0.0 $p &lt; 1$</td>
<td>0.3 $p &lt; 0.8$</td>
</tr>
<tr>
<td>Q8'</td>
<td>0.0 $p &lt; 0.9$</td>
<td>0.6 $p &lt; 0.6$</td>
<td>0.5 $p &lt; 0.7$</td>
</tr>
<tr>
<td>Q9'</td>
<td>0.5 $p &lt; 0.5$</td>
<td>0.1 $p &lt; 0.9$</td>
<td>1.6 $p &lt; 0.3$</td>
</tr>
</tbody>
</table>
the emphasized condition ($t_{130} = 3.08, p < 0.002$). As groups were given three questions to discuss, minutes 6-8 were a common time to end the first question and move to the second. Our own observation indicates that at this point, the emphasized speaker has time to observe their depiction in the Conversation Clock during the lull in conversation.

Though we continued to categorize participants as Talkative or Quiet as in past work, Table 2 indicates that the Talkative and Quiet participants reacted relatively similarly to the visualization in this setting. Changes in participation were limited to the conditions and not the individuals. For example, the number of turns consistently increased when a person was emphasized ($F_{2,48} = 16.2, p < 0.0001$) with each person taking an additional 6.2 turns per minute, up from 5.8 turns in a normal condition.
In spite of the focus on the visualization and removal of other conversational cues, our participants did not report notable pressure from the visualization or from their partner. There were changes in interaction, specifically in the length of turns. Though they provide a slight change in interaction, the distortion in the social mirror remains relatively unnoticeable.

**Remote Conversation Discussion**

Dyadic conversation can be shaped by distorted conversation visualization. With H2 we had hypothesized that emphasis would decrease the emphasized speaker’s participation. This study validates our hypothesis, though with a small aggregate difference of a few seconds.

Interestingly, as seen in Figure 7.7 the most salient moment of change consistently occurs roughly 6 minutes into conversation. This uniformly corresponds with the first major question transition. Most pairs chose to answer one question at a time and discuss it in depth before moving on. The first question often ended around 5 or 6 minutes into the session. These transitions are where balance in conversation changes. After that point in conversation, very little occurs in terms of balance. That transition between questions seems to be where most attention was paid and where the balance converge for the rest of conversation. After that point in conversation, very little changes in terms of balance (Figures 7.6 and 7.7). Similar to what is termed a midpoint transition, groups were seen to use a landmark in conversation to set the pacing for further discussion (Gersick and Hackman, 1990). However, once this first insight is made, there is very little aggregated change.

The switch to dyadic conversation impacted the types of results we could detect due to the dynamics of two person conversation (Branigan, 2006; Hare, 1981). A two person conversation has only a speaker and a listener, there is no third or fourth person to share the speaking load in conversation. The lack of contrasts between Talkative and Quiet participants can be attributed to this change. In a dyad, one cannot remain quiet and still be a part of conversation. The only distinctions made between the Talkative and Quiet in this study, is that the two groups were appropriately split into Talkative and Quiet as the Talkative consistently dominated measures of Leads, Turns, and Turn Length.

**7.6 Implications and Conclusion**

People will trust and accept a distorted image of their interaction. They will react to their distorted images. However, our results indicate this is not the primary mechanism that motivates the balance
conversation in other social mirrors. Heavily emphasized conditions saw no reported change in social pressure and only a small change in participation attributed to that distortion.

The three studies bring both practical and theoretical implications to the study of group dynamics and social computing. Validating two of our three hypotheses, we showed that people would trust a visualization over their own perception and that people would participate less when emphasized. Though we set out to demonstrate a malleable link between distortion and participation, we found there must be other mechanisms that must account for the push towards balance in conversation. We expect the dominant mechanism for change is the knowledge that one could be held accountable by the visualization. People change because they feel they are being observed and are more self conscious of their actions.

Using abstract visualization as feedback, a distorted visualization produces an effect very near the undistorted visualization. The accuracy of the visualization has a minor impact as long as the visualization remains trustworthy. At 60% additional contribution participants shifted 8% in individual contribution. Though we did not test visualizations that were so far skewed to be untrustworthy, we anticipate they would be ignored based on the ability of familiar groups to ignore the visualization and the negative reaction when the visualization was notably wrong in the pilot study.

Visualizations like the Conversation Clock articulate and help people realize what may not be readily obvious to them or to their conversation group (i.e. “I really have been talking too much,” or “I should speak up”). In the co-located group study, participants were familiar with each other and entered the study with a history of conversational patterns. In this setting, the social mirror may not be effective. Although more study is needed, we suspect that once these patterns are known by the individual and group (i.e. John talks a lot; Mary and John know that John talks a lot) and are accepted by the group, the social mirror does not provide added incentives for viewing or for modifying conversation. Once one learns the patterns from a specific group using the Conversation Clock, they may not need to keep viewing it as was the case in the second study.

Future applications of distorted mirrors would benefit from directed goals. People deferred to the visualization over their own judgment of conversation as shown in the pilot study and group study. As a motivational tool the use of distorted feedback could be an effective means to challenge people in achieving a targeted contribution goal. Work in physical therapy has been able to appropriate similar feedback to encourage recovering patients to push their muscles harder by under-reporting their performance (Brewer et al., 2005). Though the work presented here was done with participants without a specific motivation to alter their interaction, a similar approach could be taken to use social
pressure as a motivation: teaching social skills directly in conversation, conserving energy compared to one's neighbors, or increasing exercise relative to one's friends. Distorted feedback could help to drive individuals when a goal is in mind.

In the final study, we found the social mirror was most useful at a specific time. The Conversation Clock played a role in between questions at roughly 6-8 minutes into the conversation, a time when participants took a moment to break from speaking. Past work also indicates the clock is most useful when less actively engaged at the moment (not speaking or returning eye-contact). Future work might further study moments of social mirror utility to gauge if a person is watching others or specifically checking themselves. The change occurring between questions in the dyadic condition seems natural as both participants are engaged throughout the rest of the conversation. Though, an extensive categorization of the gaze direction in socially mirrored environments has yet to be done, it could show when people are most interested in feedback and potentially why.

Viewing personal social data alongside others’ personal data reveals subtle nuances of personality. Upon viewing that data, context determines how a person responds: assimilate into the crowd, to stand out as an individual, or simply know where they fit. A person can shape themselves based on what aspects they value in that context. Social mirrors have explored a subset of applications in meetings, conferences, the workplace, and therapy; they show that feedback can be used to promote characteristics that are desirable in the group. As more social interactive data becomes accessible with sensors and personal data collection, the settings for social mirror feedback only expands.

The use of these distortion strategies provides only a limited view of distortion as a whole. It investigates only the Conversation Clock and our varied set of distortion strategies in relatively short conversations. We cannot make strong claims as to how this setting affects individuals over long periods of regular use. We cannot make claims as to how other forms of distortion might affect conversation. We expect that the once a person is comfortable with the visualization, its effect would decrease unless there were a specific social reason to attend to one's representation.
Chapter 8

Conversation Clusters:
Acquiring Context from Ambiguity

As seen in previous chapters, group interfaces like social mirrors can avoid the problems of capturing content by representing interaction to promote awareness of group participation levels. Conversation Clusters pushes the context captured by these interfaces by automatically capturing the content of aural conversation. It emphasizes thematic content as graphical feedback for live review and archival. Groups and individuals benefit from creating persistent summaries of interaction. People take notes for a meeting, classroom exercise, and occasionally in casual conversation. When groups take notes on a shared interface they produce higher quality output and run more focused meetings (Olson et al., 1993). Automatically capturing content from dialog is a difficult task. The underlying system must be able to hear, understand, and summarize the major points of a conversation.

Though Conversation Clusters appears a bit different than the prior social mirrors, it was prototyped to gain access to the content of a conversation. Prior work with analyzing text demonstrates that textual analysis can be an effective manner to show meaningful stories as in Themail (Viégas et al., 2006) or highlight positivity in GroupMeter (Leshed et al., 2009). Neither of these projects model the actual content, through statistics they are able to pick out the proper words and allow the observer to make his or her own conclusions. The potential for language to play a bigger role in social mirroring is still there. The Conversation Clusters remains a prototype system to access this context. I use it here as a means to illustrate future directions in social mirrors.

Leveraging People and Computers

Understanding conversation is hard. People have evolved to quickly and easily process language in a way that computers have yet to reproduce. Computers encounter difficulties with mumbling, self correction, ambiguous references. In a laboratory environment with a system trained for a speaker, the best speech recognitions systems produce a 3% word error rate on transcripts while more common conditions incur 20-30% error rates (Munteanu et al., 2006). The word error rate is only a surface
level measuring of understanding. Properly understanding conversation requires an understanding of semantic meanings, established contexts, grammatical structures, etc. Even with a perfect transcript in the most formal language, this task is exceedingly difficult (Allen, 1995). At this time, Natural Language Processing (NLP) techniques are often limited to training and use on a specified domain. Altering the conversation topic, word choice, or context from the original training drastically reduces the automatic comprehension (Rosenfeld, 2000).

The Conversation Clusters system incorporates both speech recognition and NLP techniques to provide a contextual social mirror of conversation. Though the speech recognition is not perfect, this project demonstrates the ability to capture conversational topics by combining the strength of both humans and computers. The Conversation Clusters system attempts to model and detect the topics of conversation based on the recognized words in a conversation. The topic models can be changed and adapted by the listeners during the conversation. Corrections can be made and are incorporated into the final output: a timeline of conversation. By combining these strengths, this applied social mirror creates topically archived conversation.

**Capturing Conversation Topics**

Written language complements oral discourse as a means to exchange ideas (Ong, 2002). The predominance of literate culture in recent history has resulted in many changes in the way we remember and think as a society: written records act to accurately detail events and easily disseminate information. One of the most common artifacts of aural meetings is the transcript. They are easy to archive, store, and re-read. Computer generated transcripts using speech recognition can capture verbal exchange in a similar form. They, too, are archivable and searchable, but they are prone to errors in recognition.

With the onset of near real-time speech recognition, researchers have explored storing days and weeks of personal transcripts generated with mobile wearable microphones that continuously record. Integrated into daily life, the microphone produced an extensive database that could be stored and searched (Vemuri et al., 2006). With no specific query in mind, a cursory understanding of a transcript relies on personally skimming exact text. To date, methods such as highlighting words based on speech recognition confidence or TF-IDF scores have been used to aid in exploring these transcripts (Basu et al., 2008; Vemuri et al., 2004).

In meeting archival systems, researchers have automated detection of visual and audio cues to summarize and index video (He et al., 1999; Jaime et al., 2004). These environments are focused on
re-experiencing the raw captured media to review events. Students have benefited from similar automated capture when reviewing lectures with the eClass system (Brotherton, 2001). Other systems, such as Google’s Audio Indexing¹, have applied speech recognition to automatically create searchable indices for a video recording.

Many applications try to mitigate human involvement in computational tasks such as parsing and object recognition. More recently, popular websites like del.icio.us² utilize human knowledge to tag, label, and classify the web. Smaller groups use indices within a single document or video to easily locate useful content (Kalmkaité and Whittaker, 2008). All of these systems rely on user participation. Some, like peekaboom.org, have built games to entertain participants while capturing this information (von Ahn et al., 2006). Our goal with Conversation Clusters is to combine the computed topics with human interpretation to provide a topic labeled image of conversation.

### 8.1 Explicit Semantic Analysis and Wikipedia

We use Explicit Semantic Analysis (ESA) to extract semantically meaningful clusters of words from a text or transcript. Introduced by Gabrilovic and Markovitch, ESA analyzes a fragment of text and returns a set of semantically similar concepts from Wikipedia (the chosen source corpus) (Gabrilovich and Markovitch, 2006a). In this section, we summarize the ESA approach (see (Gabrilovich and Markovitch, 2006b) (Gabrilovich and Markovitch, 2006a) for further details).

ESA matches individual words to concepts defined in a large corpus. Each individual document defines a concept named by the title and described by the associated text. Though any large corpus might be used, Wikipedia was chosen due to the wide breadth of available concepts ranging from high profile individuals, to mathematics, to television episodes, to the articles that languish in obscurity. The ESA model builds a collection of descriptive weighted vectors based on the distribution of words in each article and throughout the corpus. Wikipedia’s index of millions of potential concepts increases the probability that nearly any topic of discussion (short of specific localities) should be detectable in our interface.

To produce a list of concepts as in Figure 8.1 the system generates a TF-IDF weighted vector for the query and maps individual words to identified Wikipedia concepts. The results for each word are weighted and merged to produce a list of the most related concepts. For our purposes, these results do not adequately describe the actual topic of discussion, though they are semantically similar.

---

¹http://labs.google.com/audi

²http://del.icio.us
The implementation provides access to a weighted list of words that triggers each result. For example, in Figure 8.1 item 1 U.S. National Monument includes (26.6, utah) (25.7, forest) and (13.6, preserv) as trigger words, whereas item 2 Utah Lake provides (60.4, utah) and (13.6, preserv). These trigger words vary more often with larger queries, highlighting multiple concepts that might be represented in the initial text. At the same time, ESA links these clusters of words to a common theme, providing a meaningful approximation of the current topic.

Our algorithm, built on top of this source, attempts to connect the topics at each moment in time into themes that develop and change as the transcript progresses. Other word clustering algorithms mentioned above could replace ESA in our implementation; we found the access to human specified semantic relations of words in Wikipedia proved an effective tool in extracting related words from an error laden speech transcript.

8.2 Clustering Conversation

The goal of Conversation Clusters is to capture the “gist” of conversation. This content should be usable in real-time and accessible afterward. As in Leshed’s work, a constant computational awareness of content would allow people change their interaction based on a mirrored depiction. Additionally the content augmented social mirrors have a stronger role to play as archival tools.

Conversation Clusters approaches dialog as a stream of ever changing topical threads. Rather than focusing on reproducing an error free word-for-word transcript, we detect general topics from the words that we can reliably recognize. A dynamic visualization on a shared public tabletop enables participants to see the most recent discussion topics as clusters of words grouped together (Figure
Figure 8.2: Captured conversation clusters on the table in real time. The clusters reveal the current model of conversation as the computer sees it. Individual people can correct this model directly. Users can also switch to a corresponding timeline visualization to explore the birth and death of topics. With Conversation Clusters, we develop a representation of conversation that captures the context of the moment.

Clusters

The basic cluster visualization depicts current and recent topics (Figure 8.2). The visualization depicts the computer's topic model as an ambient display. Each cluster represents a detected topic while free words hint at possible other topics. The entire visualization can be manipulated via Mitsubishi's DiamondTouch table to alter the underlying topic model.

As with the other social mirrors, the visualization occupies public conversation space. Since this visualization utilizes actual words and content, the visualization attempts to mitigate distraction. The words in the visualization move slowly; changes to topics are accomplished with a gradual fade or by reusing elements. A word replaced in a cluster will fade from one word to another rather than add one word and subtract another. However, clusters themselves are reused only in cases where the new topic is sufficiently close to the old (we discuss the distance metric in a later section). Emerging topics in conversation that cannot be paired appear as a new cluster. As a group moves into new topic areas, the shared display demonstrates the changing focus of the meeting with multiple new topic clusters.
Words unrelated to the highest rated clusters move freely and unbound in the visualization. These words may be from a previous topic that is no longer being discussed or an emerging topic that may soon appear on the table. With the interactive tabletop, participants can increase the strength of a link between the lesser topics and current clusters by dragging the floating words into clusters. With no interaction, the words simply demonstrate the changing pattern of topics. As speech recognition and topic detection algorithms are not perfect, erroneous words and glaring errors do appear and may be removed by pressing a full hand to the word. In this manner, people present can influence topic formation and topic removal.

**Thread History**

The thread history visualization shows a historical overview of salient topics in discourse over time. As seen in Figure 8.3, threads of conversation begin, merge, separate, and end along a timeline from left to right. The computer diagrams history by creating and displaying this model of conversation evolution.

![Diagram of thread history visualization](image)

**Figure 8.3:** The thread history visualization demonstrates the progression and change of conversational topics.

In this historical view, multiple threads appear in parallel. This does not mean that four or five topics are being separately addressed at once, but demonstrates related topics that are not clustered together. For example, a discussion about car repair might have a thread about cars and another about the financial means of payment. In the context of the conversation they are related so they appear simultaneously. However, the words that generate these threads do not sufficiently overlap to form a single cluster or thread. Additional parallel threads appear as the transcript grows in size. The thread history visualization examines the relation of clusters over time, and the window of potential related clusters grows proportionally. The interface allows for interactive exploration. Longer conversations
are viewed at a higher level; participants can zoom in to see a more detailed view. They can also switch between the timeline and cluster modes of the table visualization.

The timeline visualization creates an abstracted view summary to aid recall by filtering words upfront. The full dialog cannot be recreated from the visualization. Threads of conversation use individual salient words to characterize the evolving topic. The selected words allow viewers to infer the general meaning but provide privacy for the individual speakers.

8.3 How to Trace Themes

Conversation Clusters provides an interface to the constantly evolving topics of conversation. The mirror is a visual artifact that can act as a window into that content. This section shows how Conversation Clusters generates that model. Like much of the related work statistically representing language, the initial classification with ESA only provides a summary relevant to the queried text. Though this type of overview can be useful, tracing themes requires the algorithm to connect these patterns over time.

The transcript is first segmented into arbitrary segments of time. After segmenting the text, each segment becomes an ESA query consisting of the top ten results. We keep only the concept weight, the trigger words, and the word weights. We then group each query into a 'bin' according to their chronology. In a document this might be by paragraph or section; in a continuous text stream this might be by a predetermined time interval.

\[
\begin{align*}
\text{Topic}_1 & : (w_1, a_1) \quad (w_2, a_2) \quad \ldots \quad (w_x, a_x) \\
\text{Topic}_2 & : (v_1, b_1) \quad (v_2, b_2) \quad \ldots \quad (v_x, b_x) \\
S(T_1, T_2) & = \frac{\sum_{i=1}^{\min\{x, y\}} (w_i + v_i) + \sum_{i=1}^{\min\{x, y\}} (w_i + v_i)}{\sum_{i=1}^{\min\{x, y\}} (w_i + v_i)} \\
d(T_1, T_2) & = S(T_1, T_2)^{-1}
\end{align*}
\]

**Figure 8.4: To calculate the similarity and distance of two topics**

The bins offer a balance between capturing the detail in each segment while still allowing the ability to trace topics. However, many potential topics are redundant within a single query and between similar queries in the same bin. We mitigate this repetition by combining topics within a bin. The distance metric in Figure 8.4 emphasizes co-occurrence of the topmost words in two topics. The distance ranges from 0 (no words matching) to 1 (all words match with the same weights). Combining the similar
topics is a matter of a linear combination. The resulting new topic words and weights is a weighted sum of the original topics (Figure 8.5). The final word ranking determines the topic’s description and favors the highly weighted words.

\[ T_{new} = \max(T_1, T_2) \]
\[ u = \frac{T_1 w_i + T_2 v_j}{T_1 + T_2}, \text{ where } a_i = b_j \]

\[
\text{Topic new } T_{new} : (u_1, c_1) (u_2, c_2) \ldots (u_z, c_z)
\]

**Figure 8.5:** Recalculating topic and words weights when combining two topics

![Figure 8.5: Recalculating topic and words weights when combining two topics](image)

Redundancy elimination continues until no two remaining topics are closer than a threshold value. We found .35 to be an adequate starting threshold; however, if the bins cover larger time periods and thus more queries, higher thresholds are needed. When the cutoff is reached, the bin contains a selection of distinct topics. The bins are compared with neighboring bins using the same distance formula discussed above (Figure 8.4). A connection between topics, as seen in Figure 8.6, indicates the latter topic is a continuation of topics seen earlier in the text. Some themes stretch throughout the entire timeline while others indicate a chronologically localized topic.

Once the themes have been extracted, the final step is to choose words from the source text to depict the topics over the theme length. Starting from the earliest topic in the theme, the theme displays the
highest weighted topic word by default. Later topics in the theme are constrained to show the highest weighted word that has yet to appear in the theme (Figure 8.7). Any point on the timeline can be explored by touching a word to reveal the top set of words, but the single word choice demonstrates the main theme as it begins and how it refines itself over time.

**Figure 8.7:** Choosing the words to represent a thread of conversation: the highest ranked word that is current and unused

### 8.4 Applications and Limitations

In this dissertation, Conversation Clusters was by far the most intricate and complicated social mirror to construct and run. For a group of individuals, the system required a computer per speaker for speech recognition, another for the visualization, and a final one to handle the clustering. The base level accuracy of speech recognition could be improved by incorporating time-consuming training beforehand. While certainly not a practical setup for most casual groups, the Conversation Clusters interface demonstrates a proof of concept for capturing context from noisy signals.

From the beginning, I adopted an artistic approach with the intention of later appropriating methods from Conversation Clusters into other social mirrors. Thus the interface was chosen as much for it's aesthetic and soothing value as it was the functional values of demonstrating clusters. Though the clustering interface came very close to my sketched prototypes, the thread history visualization never quite reached it’s conception. The image shown in Figure 8.8 depicts what I really wanted to achieve with this prototype, an easily readable summary of the rising and falling topics of conversation.
At present, Conversation Clusters allows for the extraction of topics in near real-time during a conversation, however it can quickly capture an overview of content from any stream of text. More importantly, the topics are accessible to an interface designer. Thus, the topics can be incorporated as visual markers in conversation for when topics change radically, as summaries of multiple ongoing online discussions, or applied to a document to understand its structure. Applied to a large corpus of audio, this technique enables a way to quickly skim archived audio. A journalist, an anthropologist, or a linguist with multiple hours of recordings could search for specific items he or she remembered. With a visual depiction, they could focus their review to most pertinent sections based on theme.

The basis of our algorithm is a statistical approach based on a finite set of topics. The system cannot easily detect location specific topics unless they appear in the corpus. New concepts and potential topics will also remain obscure until the corpus is reanalyzed. However, the purpose of this visualization is to provide an accessible and persistent visual cue into past interaction. While an exact match might not occur, the broad range of concepts in Wikipedia ensures that there is something semantically close to extract a topic.

The Conversation Clusters real value lies in its underlying representation of conversation. The other visualizations in this dissertation are devoid of content from the conversation itself. Simply augmenting
the Conversation Clock or Conversation Votes with a Conversation Clusters back end could provide a significant amount of context into their real-time display and improve their archival value. The Conversation Clock could be summarized in rings, allowing an individual to see the conversational patterns alongside the content. One could also see if a person dominating the conversation was speaking on topic or derailing the conversation. In its present form, the Conversation Clock does not offer much value for an individual not present in the conversation. Similarly with Conversation Votes, capturing the voting feedback could highlight important content along with the speaker — offering a better summary of the most salient moments in conversation.

Conversation Clusters demonstrates a method that is effective at capturing that content. For reasons of time and resources, we did not pursue further integration into the social mirrors in order to more fully explore the results we had already seen. I still see the incorporation of context into these visualizations as a valuable path to be explored in future studies.
Chapter 9

Conclusion

This dissertation may come to a conclusion but there are still questions to answer. Below I summarize my findings before continuing to further discuss the limitations of my own work and future directions that it may yet explore.

9.1 Contributions

In this dissertation, I have presented the foundation of social mirrors as a theoretical construct. I have illustrated a variety of social mirrors of my own creation and shown how people respond in the presence of them. I believe this work addresses fundamental aspects of face-to-face interaction as we integrate more technology into our daily lives and encounter increasingly various modes of interaction online. Below, I will revisit each of my initially stated contributions to summarize the major results of my work.

1. The social mirror provides a theoretical framework to describe new interfaces. The social mirror as defined in Chapter 3 sets out a description for self-evaluation interfaces. This type of interface is meant for the individual to better understand his or her situation and influence that individual in the moment. It is a heads-up display for everyday life.

My own studies were limited to verbal speech, but the social mirror construct can be applied to other scenarios to demonstrate visualization as an instigator of self-motivation. Consider applying social mirrors to environments that traditionally suffer from lack of feedback, such as improving teleconferencing and augmenting text-based conversations (email, instant messaging, chat). Consider the capture of sentiment, mood, and aggressively from word choice and inflection to illustrate the tone of conversation. The mirrors can be extended to be much more explicit in what they show and to be applied in new environments.
However, the social mirror has more potential in creating new cues and making them available for social critique. Expanding from conversation, it would be a small step to apply social mirrors to energy consumption, increasing physical activity, reducing wasted time on a computer, or revealing the amount of work that went into a document. Each of these scenarios can benefit people by making a person responsive to social pressures.

2. Abstract audio visualization and aesthetic design can be leveraged to convey information in visually active environments. With each visualization presented in this dissertation, the visualization’s context and purpose influenced the design. Particularly in collocated conversations, the visual signals of the interface compete for attention with traditional conversation signals. I have shown visualizations that distill unseen patterns of conversation, turns, interruptions, domination and made them into active signals in conversation.

My design contribution offers a history of design choices. For each visualization, numerous potential designs influenced the final product. As with any iterative process, the final product would not exist without the steps taken along the way. My final designs represent the culmination of my work on these projects, but not a finality to the future development of social mirrors and conversation interfaces.

3. Real-time visualization can empower individuals in conversation. In Chapters 4–7 I have shown social mirrors in casual conversations, group project meetings, debates, and classrooms; in each setting real-time self evaluation was able to improve interaction for some individuals (discussed in my next point). In each case we saw an instance where the underrepresented better asserted themselves in conversation. In general, the social mirror produced a “balancing” effect of speech contribution causing the talkative to speak less and the quiet to speak more. The Conversation Clock did this, Conversation Votes did this, and — to some extent — the Fragmented Social Mirror did this.

Empowerment was not limited to just a balancing of speech contribution. Conversation Votes and Fragmented Social Mirror demonstrated that there was a silent group that wanted to speak out. In Conversation Votes, the individuals who were not satisfied that opinions were as diverse as they should be spoke more and were then more satisfied with the discussion as a whole. In the Fragmented Social Mirror, students acknowledge that they have questions to ask, but they do not ask them. The real-time channel allowed them to get involved and initiate interaction.

However, with Chapter 7 we see a particularly interesting effect: the real-time feedback need not
be fully accurate to see the balancing effects. With a notably distorted visualization, the balancing
effects still occurred. Though, I showed that there is some effect that the distortion causes, it is
small and limited. It seems the bulk of the effect of social mirrors is in fact due to the perceived
sense of accountability rather than the details shown in the visualization.

4. **Simple feedback from the crowd effectively conveys group information.** Shown in Chapter

5 I demonstrated that the used of anonymous feedback could convey group information. With
Conversation Votes this was seen as a sense of approval or agreement. Participants reported
moving on to other topics when they saw that others agreed with them. Or, in some cases,
they noted that the other listeners were in agreement and chose to end their turn. With the
Fragmented Social Mirror we see this at a much larger scale. We see that the hundred or so
students would not speak up. However when they ask their questions, the lecturer can see how
well the class is understanding the material. The lecturer now has the ability to gauge the class in
their understanding in a way that was not possible and it opens a new dialog of communication.

I chose to use anonymous feedback as a way to encourage more individuals to break that barrier
from silence. As mentioned in the limitations section earlier, this contribution should be later
tempered with a comparison between anonymous feedback and non-anonymous feedback in
future work.

5. **Social Mirrors produce meaningful visual archival of audio that can be glanced over.** All
of the social mirrors in this framework capture conversation and demonstrate visualizations.
However, most of these visualizations are not appropriate for long term archival. There is no sense
of context or history preserved in the visualization. Chapter 8 presented Conversation Clusters as
the thematic progression of conversation. It demonstrated a proof of concept showing that the
real-time context of conversation could be detected, modeled, and improved by individuals in the
conversation. The same technique can be applied to existing social mirrors to make them better
archival references or to be extended in its own right as a new tool.

No work is ever truly complete, there are certainly avenues I desire to pursue in the vein of social
mirrors and other shortcomings. I highlighted a number of limitations along the way: limited to small
groups, small samples, certain populations. There are larger limitations that prevent generalizing
across all conversations and all domains. The subset of studies I pursued were chosen address the
effects of real-time feedback on conversation. Some of the following studies were eliminated due to
technical or time limitations. I still see them as viable options to continue investigations into social
mirroring.
Longitudinal Effect of Social Mirrors

Each social mirror was created and studied in a single context (with the exception of the Conversation Clock). They were built to test an idea, that idea was tested, and then I moved on to test another aspect of design.

In the initial studies, the experiments were run first with a no-visualization condition, followed by a visualization condition, followed by a final no-visualization condition. The rationale was that the visualization would potentially make people more aware of their interaction and change it for future conversations. They might remember how they had previously appeared and tailor their interaction based on that perception. The initial studies with Conversation Clock and Conversation Votes did not show such results, and it was never a primary focus in my work. However, there is still reason to explore the social mirror in longer term changes. Much like therapy or practice, it may take time and repeated exposure. The questions still remain as to whether the effects can "stick" and have an effect without the social mirror present. Do people acclimate to the social mirror? Do they continue to have an effect after weeks of use? Are the effects long lasting if a group had been using the mirror for several weeks or months? Specifically for the feedback, do people continue to use feedback regularly and does it remain useful?

Along similar lines, the longitudinal studies should examine social mirror use in varied real-world settings. I would not expect identical effects in a business meeting as I would in a brainstorming session. These are the types of domains that can be fully explored only through long term evaluation.

Storytelling

In addition to longitudinal effects, I am interested in the act of storytelling based on archived visualizations. Storytelling is using the visualization as an aid to recall memories and conversation after the fact. Viégas et. al explored similar effects with email histories in (Viégas et al., 2004a) and (Viégas et al., 2006).

Storytelling is situated in stories that the viewer is interested in recalling. With email, there is a broad range of time and topics involved, allowing a visualization to cut through and identify the most salient interactions amongst them. With the social mirrors, they were not yet so ubiquitous. To draw upon the patterns of conversations and identify stories, the participant should be able to see past and notable conversations that they were familiar with. Conversations with loved ones, daily patterns with friends,
routine interactions at the coffee shop, etc. While Mathur did apply conversation visualization to conversations in Skype\(^1\) (Mathur, 2009), it remained a limited conversation space.

To truly explore storytelling in visualized conversation, I need constant access to conversations, knowledge of participants, and quality audio input. For various reasons, including time and infrastructure, this domain remains unexplored for future work.

**Anonymity versus Non-Anonymity**

Chapter 5 explored the use of anonymous feedback to augment conversations. I chose to explore anonymous feedback based on related literature that indicated people would be more likely to voice their opinions, particularly if there is a hierarchical relationship such as a boss and employee. Though I did see participation in my system, effective in both Conversation Votes and Fragmented Social Mirror, there are no contrasts with non-anonymous feedback. I am left wondering if the mere opportunity to provide feedback would be as effective if names and identities remained attached.

The question of anonymity is particularly relevant for the classroom setting of Chapter 6. Identity may help stem the off-topic conversation, but how does the feedback change in character as a result of adding identity. Do the students become as reserved as they were without the Fragmented Social Mirror? Does identity limit the students who use the Fragmented Social Mirror to those that already speak in class? How much of the increased activity we saw as a result of the Fragmented Social Mirror is due to opening a public backchannel that can be accessed at any time versus that same backchannel tied to identity.

This is a question that I did not answer, but I find it to be an interesting one and a particularly important one to pursue in order to inform classroom systems.

**Real-Time versus Post-Meeting**

All projects explored in this work looked at visualization as a real-time motivator. This is related to the motivation for change. As discussed in Chapter 4, the social mirror is grounded in the interplay of signals and knowledge of what is seen. People react to social cues in the moment. A post-meeting visualization in inherently unseen until after the conversation has ended.

\(^1\)A popular internet telephony program
I investigated how people sought to use the visualization to empower themselves and alter themselves in conversation — according to their own interpretation of how they should appear. Providing the visualization afterwards seemed better suited to situations where a desired outcome is predetermined.

Though I did not investigate it in my own work, I see this comparison as a potentially useful study to investigate. Particularly for goal oriented or therapeutic visualizations, post-meeting visualizations could allow an individual to confirm or refute personal progress.

**Applications for Social Mirrors**

As with the longitudinal studies, the social mirrors would benefit from being placed into focused domains. Though, I might apply them in multiple areas such as meetings, brainstorming sessions, classroom, I would like to focus on one area that we began to pursue — teaching social skills. This work is actually still in progress due to many encountered many snags along the way.

The Conversation Clock could act as a therapeutic tool for people diagnosed with a variety of social language deficits including Asperger’s Syndrome (AS), Non-verbal Learning Disorder (NVLD), High-Functioning Autism (HFA), and Pervasive Developmental Delay (PDD-NOS). Children and Young adults diagnosed with this set of disorders are prone to ignore social protocol and miss cues in conversation that are obvious to others (Attwood, 2006). As a result, they often appear tactless due to an impaired ability to understand another’s mental or emotional state (Baron-Cohen, 1997). People with these diagnoses often seek help through therapy to better develop their social skills. A diagnosed individual might interrupt speakers, dominate conversations, and misread (or not read) cues from the listeners. Additionally, their comments might be seen as overly critical and inappropriate.

Traditional therapy for people with social language deficits consists of explicitly teaching rules of interaction, practicing skills in related scenarios, and reviewing their use during a video based social autopsy. There are few opportunities in treatment that allow for the immediate modification of behavior. The social mirror provides feedback that will enable diagnosed individuals to evaluate and change their interaction during their scenario.

Many children, diagnosed on the spectrum, see a reduction in characteristic symptoms with therapy (Piven et al., 1996). A review of social skills therapy shows that complex social behavior should be made concrete through either visual or tangible activities (Krasny et al., 2003). The use of computers and other technology has been found to both effective at helping children while also being appealing to the child (Goldsmith and LeBlanc, 2004).
The Conversation Clock fits perfectly in this domain, as it demonstrates many of the types of social cues that people diagnosed with these disorders have problems detecting. For each participant, we assess a broad range of indicators on language skills, social communication abilities, and learning style preferences. These initial measurements provide insight that we can use to target future populations.

9.2 Final Words

Social Mirrors allow us to better understand ourselves and our interactions. They capture a different perspective: a view others see that we do not. By our own initiative, we shape our interactions to better reflect the self we want to portray. This work demonstrates visualizations can become natural extensions of our view of self. With the proper design consideration and perspective, the Social Mirror is both a part of the environment and a part of the conversation. Mirrors allow us to look at ourselves and change, Social Mirrors are expanding the domain of what we can examine and how we can change.
### Appendix A

**Conversation Votes:**

**Debate Questions**

**Session A: Should smoking be banned in public places?**

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientists agree that smoking is dangerous. Tobacco smoke can cause</td>
<td>Society accepts that adults can decide to harm themselves to some</td>
</tr>
<tr>
<td>cancer, strokes and heart disease. Smoking does not just harm the</td>
<td>extent, so long as they do not harm others. This is why the</td>
</tr>
<tr>
<td>smoker $\Box$ it also harms people nearby, who breathe in the</td>
<td>proposition is not arguing that people should be banned from</td>
</tr>
<tr>
<td>smoke. Smokers choose to smoke, but people nearby do not choose to</td>
<td>smoking in private. Passive smokers do choose to breathe in other</td>
</tr>
<tr>
<td>smoke passively.</td>
<td>people's smoke. If they do not want to smoke passively, they do</td>
</tr>
<tr>
<td></td>
<td>not need to go to places where smoking is allowed. There is</td>
</tr>
<tr>
<td></td>
<td>therefore no reason to ban smoking in public.</td>
</tr>
</tbody>
</table>

The opposition is wrong to say that people choose to smoke passively. In many places, there are no non-smoking bars or restaurants. Unless people refuse to go out with friends, they cannot avoid passive smoking. People who work in smoky workplaces (e.g. bars) often do not freely choose this sometimes no other jobs are available.

If enough people want to go to non-smoking bars, companies will set up non-smoking bars. If there are no non-smoking bars, this suggests that very few people want them. Some people are quite happy to work in smoky places. A complete ban is not necessary to protect workers $\Box$ ventilation fans can remove most smoke.
A ban would encourage smokers to smoke less or give up. If smoking was banned in public places, it would no longer be a social activity. Instead, smokers would have to leave their friends inside and go outside to smoke. This would be particularly unpleasant when it is cold/wet.

People will not smoke more at home. Smokers need to maintain a certain level of nicotine in their blood to remain content. A ban on smoking in public would force them to smoke less while at work. Over time, this would lower the level of nicotine they need to feel content.

It is legal to smoke tobacco, so governments have no right to try to make people stop. It is therefore wrong to argue that a ban on public smoking should be introduced to encourage people to give up. Smokers fund their own healthcare through the high taxes they pay on tobacco.

Banning smoking in public will encourage people to smoke more at home. This will harm other people in their house, particularly children. This is important, since children are not old enough to choose freely to smoke passively.

A ban on smoking in public places would drive many bars, pubs and clubs out of business. Smokers would not go to these places.

There have been few problems with bans where they have been introduced. Heavy fines put off companies from allowing people to smoke. A survey for the Scottish Executive found that 99.4% of premises were observing the ban three months after it was introduced.

It would be impossible to police this ban in many public places. Small workplaces will often ignore the ban and are unlikely to be caught. Staff who do not smoke are unlikely to report smokers, in case their colleagues work out who told the authorities.
Session B: Should governments introduce identity cards and require each citizen to carry one?

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each man, woman or child has a unique identity</td>
<td>Identity cards are an attack by governments on privacy and basic freedoms. Why should we be required to prove who we are all the time?</td>
</tr>
<tr>
<td>- their body, face, mind and personality are not quite like anyone else’s. For this reason knowing who you are and being able to prove it are basic human rights.</td>
<td>Don’t we have the right to be left alone by the government?</td>
</tr>
<tr>
<td>Introducing identity cards would help prevent terrorism. If police officers and immigration officials can check up on people behaving oddly, then they will be able to prevent bombings and other terrorist attacks.</td>
<td>Identity cards will not stop terrorism. Many of those guilty of terrorist attacks such as 9/11 and the Madrid bombing were in the country completely legally, and were unknown to the police. The danger is that bringing in cards will create a phony sense of security.</td>
</tr>
<tr>
<td>Identity cards will help reduce crime and illegal immigration. Those who have nothing to hide have nothing to fear if a policeman asks them to identity themselves. But criminals and illegal migrants will find it harder to go about their business.</td>
<td>Identity cards will do nothing to cut crime and illegal immigration. Even if the cards could not be faked, criminals and illegal migrants will carry on hiding from authority. If they do not come to the attention of the police, then they can’t be asked to identify themselves. And if they are stopped and don’t show a card, what will happen to them? Can’t they just say that they have lost or forgotten it?</td>
</tr>
<tr>
<td>Cards will tackle new types of crime like fraud and identity theft. Banks, individuals and companies lose billions every year to people who pretend to be someone else. The government and taxpayers lose billions every year in fake welfare claims.</td>
<td>Cards won’t stop fraud or identity theft. Such crimes are now carried out online without any need to show yourself in person. Instead these cards put our very identities at risk. Because they contain so much personal information they will replace many different types of existing identification.</td>
</tr>
</tbody>
</table>
Identity cards will be impossible to fake. Using biometric date, the identity could be positively identified by an official with a biometric reader.

Biometric science is still very unreliable. Readers are not very accurate and so lots of innocent people will end up in trouble with the authorities. We would have to be 100% sure the central database of everyone’s biometric details can’t be hacked into and changed.

Life will be easier for all of us if we have identity cards. At the moment we all carry many different cards and papers in fat wallets or purses. But one smart card can carry all this information and more.

Identity cards will make life harder for all of us. We will be expected to carry a card at all times - even on the beach, or while exercising, or in a nightclub. And if we can’t show one when asked, we will be punished.
### Session C: Should the age at which people gain the right extend to 16?

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 year olds are mature enough to make important decisions such as voting. Their bodies are fully adult, they have been educated for at least 10 years, and most of them have some experience of work as well as school. There is no magic difference between 16 and 18.</td>
<td>16 year olds are not mature enough. The large majority still live at home and go to school. They may have adult bodies, but their minds are still those of children who have to be protected. By 18 they have become much more independent and are able to make their own way in the world.</td>
</tr>
<tr>
<td>16 year olds already have other rights. For example, in many countries they have the right to leave school and leave home, and the rights to have sex, marry and have children. It is not reasonable to have different ages for different rights.</td>
<td>It makes sense for different rights to be gained at different times as young people mature and get used to more responsibility. Because voting is so important, it should be one of the last rights to be gained.</td>
</tr>
<tr>
<td>It is unfair to have taxation without representation. Many 16 year olds work and pay tax on their earnings. At present they are not allowed to have a say in how the government spends their money, nor in how much should be collected from them in taxes.</td>
<td>We trust parents to cast votes after thinking about the interests of their families. And there are other ways for young people to have a say - they can write to elected representatives and newspapers, sign petitions, speak at public meetings, join youth parliaments, etc.</td>
</tr>
<tr>
<td>It is important in a democracy to include as wide a range of opinion as possible. The idea of representation by elected officials means that as many different views as possible should be heard at election time. Teenagers have their own interests and views which are currently not represented.</td>
<td>Everyone would agree that there has to be a minimum voting age. Most people believe that the line should be drawn at 18 rather than 16. Although some 16 year olds may be mature enough to vote, most have not yet formed political views of their own yet. On average, young people are much more likely to be ready for the responsibility of voting at 18.</td>
</tr>
</tbody>
</table>
There is a problem of apathy in many western countries, with low turnouts at elections. Although young people are taught citizenship or civics at school, they don’t get a chance to put this knowledge into practice for several years. Is it surprising that they lose interest in public affairs during this time?

At the moment, 18-25 year olds are the least likely to cast a vote at election time. Youth membership of political parties is falling. Most people don’t vote because they think the election system is unfair, or because they don’t trust any of the political parties on offer - lowering the voting age won’t solve these problems.

Giving the vote to teenagers would force politicians to take them seriously. Policies on education (e.g. student loans) would have to take their views and interests into account for the first time. 16 year olds today are well-educated and media-savvy, so they can express informed opinions.

It would be dangerous to give young people the vote. They might use it in foolish ways. They might put extremists into power or vote without thinking on single issues (e.g. making drugs legal, free university places, cheap beer!).
### Conversation Votes: Extended Tables

<table>
<thead>
<tr>
<th>Measure</th>
<th>F value</th>
<th>p value</th>
<th>Category</th>
<th>Pre-Vis</th>
<th>With Vis</th>
<th>Post-Vis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>µ</td>
<td>σ</td>
<td>µ</td>
</tr>
<tr>
<td>Leads</td>
<td>$F_{(2,21,3)} = 5.42$</td>
<td>$p &lt; 0.012$</td>
<td>Talkative</td>
<td>19.8</td>
<td>5.1</td>
<td>17.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>7.4</td>
<td>3.1</td>
<td>9.5</td>
</tr>
<tr>
<td>Turns</td>
<td>$F_{(2,20,3)} = 2.32$</td>
<td>$p &lt; 0.13$</td>
<td>Talkative</td>
<td>0.92</td>
<td>0.37</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>0.38</td>
<td>0.12</td>
<td>0.46</td>
</tr>
<tr>
<td>Turn Length</td>
<td>$F_{(2,24,1)} = 0.21$</td>
<td>$p &lt; 0.82$</td>
<td>Talkative</td>
<td>22.7</td>
<td>10.1</td>
<td>21.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>16.6</td>
<td>6.2</td>
<td>18.3</td>
</tr>
<tr>
<td>Votes</td>
<td>$F_{(2,19,1)} = 1.08$</td>
<td>$p &lt; 0.36$</td>
<td>Talkative</td>
<td>0.27</td>
<td>0.34</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>0.27</td>
<td>0.45</td>
<td>0.26</td>
</tr>
<tr>
<td>V. Effect</td>
<td>$F_{(2,38,0)} = 0.30$</td>
<td>$p &lt; 0.75$</td>
<td>Talkative</td>
<td>0.09</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>0.09</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td>Represented</td>
<td>$F_{(2,22,6)} = 0.14$</td>
<td>$p &lt; 0.88$</td>
<td>Talkative</td>
<td>5.7</td>
<td>1.2</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>5.4</td>
<td>1.3</td>
<td>5.9</td>
</tr>
<tr>
<td>Comfortable</td>
<td>$F_{(2,22,6)} = 0.64$</td>
<td>$p &lt; 0.54$</td>
<td>Talkative</td>
<td>6.6</td>
<td>0.5</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>6.6</td>
<td>0.6</td>
<td>6.5</td>
</tr>
<tr>
<td>Opinion</td>
<td>$F_{(2,22,5)} = 0.91$</td>
<td>$p &lt; 0.42$</td>
<td>Talkative</td>
<td>1.8</td>
<td>0.9</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>2.8</td>
<td>1.5</td>
<td>2.8</td>
</tr>
<tr>
<td>Error Est</td>
<td>$F_{(2,24,9)} = 0.23$</td>
<td>$p &lt; 0.80$</td>
<td>Talkative</td>
<td>0.37</td>
<td>0.22</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>0.42</td>
<td>0.17</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Table B.1: This table shows differences of the visualization’s effects on Talkative and Quiet participants.
<table>
<thead>
<tr>
<th>Measure</th>
<th>F value</th>
<th>p value</th>
<th>Category</th>
<th>Pre-Vis</th>
<th>With Vis</th>
<th>Post-Vis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(\mu)</td>
<td>(\sigma)</td>
<td>(\mu)</td>
</tr>
<tr>
<td>Leads</td>
<td>(F_{(2,21.3)} = 5.35)</td>
<td>(p &lt; 0.013)</td>
<td>Active</td>
<td>14.9</td>
<td>6.4</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Less Active</td>
<td>12.1</td>
<td>8.6</td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.62</td>
<td>0.30</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.68</td>
<td>0.47</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21.3</td>
<td>9.0</td>
<td>19.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18.0</td>
<td>8.6</td>
<td>20.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Active</td>
<td>0.53</td>
<td>0.42</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Less Active</td>
<td>0.01</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.10</td>
<td>0.07</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.07</td>
<td>0.04</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Active</td>
<td>4.9</td>
<td>1.3</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Less Active</td>
<td>6.1</td>
<td>0.8</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.5</td>
<td>0.5</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.5</td>
<td>0.6</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Active</td>
<td>2.2</td>
<td>1.4</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Less Active</td>
<td>2.4</td>
<td>1.1</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.38</td>
<td>0.17</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Less Active</td>
<td>0.41</td>
<td>0.21</td>
<td>0.31</td>
</tr>
</tbody>
</table>

**Table B.2:** This table shows differences of the visualization’s effects on Active Voters and Less Active Voters.

<table>
<thead>
<tr>
<th>Measure</th>
<th>F value</th>
<th>p value</th>
<th>Category</th>
<th>Pre-Vis</th>
<th>With Vis</th>
<th>Post-Vis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(\mu)</td>
<td>(\sigma)</td>
<td>(\mu)</td>
</tr>
<tr>
<td>Leads</td>
<td>(F_{(2,21.3)} = 5.35)</td>
<td>(p &lt; 0.013)</td>
<td>Heavily</td>
<td>13.7</td>
<td>5.7</td>
<td>12.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lightly</td>
<td>13.3</td>
<td>9.3</td>
<td>14.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.65</td>
<td>0.45</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.65</td>
<td>0.33</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21.1</td>
<td>8.2</td>
<td>18.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18.2</td>
<td>9.4</td>
<td>21.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.37</td>
<td>0.44</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.17</td>
<td>0.32</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
<td>0.05</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Heavily</td>
<td>5.0</td>
<td>1.3</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lightly</td>
<td>6.0</td>
<td>1.0</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.5</td>
<td>0.5</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.5</td>
<td>0.6</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Heavily</td>
<td>2.3</td>
<td>1.4</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lightly</td>
<td>2.3</td>
<td>1.2</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.35</td>
<td>0.17</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.43</td>
<td>0.21</td>
<td>0.33</td>
</tr>
</tbody>
</table>

**Table B.3:** This table shows differences of the visualization’s effects on Heavily Supported and Lightly Supported participants.
APPENDIX C

Fragmented Social Mirror:
Dialogs

This section documents the observed behavior in the classroom before and after the introduction of the Fragmented Social Mirror.

C.1 Control Observation

Session 1

<table>
<thead>
<tr>
<th>Instructor</th>
<th>Can we pick the smallest integer in this set?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>yes [mumbled, head-nods]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instructor</th>
<th>$n$ is divisible by what?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>$r$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instructor</th>
<th>So what is $a - b$?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>$r$</td>
</tr>
</tbody>
</table>

Session 2

<table>
<thead>
<tr>
<th>Instructor</th>
<th>How many know what a Corollary is?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>[hands raised]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instructor</th>
<th>What happens if the two numbers are reversed? [swapped]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>[no noted response]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instructor</th>
<th>Does it stop here?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>No.</td>
</tr>
<tr>
<td>Instructor</td>
<td>Right, technically the algorithm keeps going before we return.</td>
</tr>
</tbody>
</table>

Session 3
Instructor: How many people know what this $\varnothing$ is?
Class: Empty set.

Instructor: What is the cardinality of $Q$?
Class: \(\text{three.oldstyle}\)

Instructor: \((\text{one.fitted, two.fitted, three.fitted})\) is a subset of \((\text{one.fitted, two.fitted, three.fitted})\)?
Class: [head nods]

Instructor: Yeah, it's true.

Instructor: The empty set is a subset of any random set?
Class: Yes/No/[head nods]

### C.2 Experimental Observation

**Session 4**

Instructor: Looks like the screen could be brighter [Screen is dark, hard to read]

\[\text{\xmark}\]

: [Screen is made lighter]

\[\text{\cmark}\]

: [An assortment of positive and negative feedback appears as new people join and test their connection. No connection to the content is noted]

Instructor: \(A\) is contained in \(B\), \(B\) is contained in \(C\) - would someone like to say where I'm headed?
Class: [mumbled] \(A\) contained in \(C\)

(2) \[\text{\cmark}\]

: Test

\[\text{\cmark}\]

: t-t-t- TEST MESSAGE

Instructor: Therefore, we can conclude the sets are equal.

(10) \[\text{\cmark}\]

Instructor: [Anecdote about difficult problems in a Qualifying exam, but moved on]

: What was the claim that they were laughing at?

(2) \[\text{\cmark}\]

Instructor: I have no idea, something the guy thought they ought to be able to prove, but he was wrong.
Instructor: So what do we get when we cross two sets?
Class: ordered pair

Instructor: [talking about mixed conventions on numbering proofs as a way to save space]
Class: Are we allowed to use numbering at all?

Instructor: Yes, go ahead.

Session 5

Instructor: [Explaining why the Quiz has yet to be returned]

Instructor: Discussing the Karatsuba Algorithm
Class: <What is Karatsuba?>
Instructor: Karatsuba is the guy who invented it, Anatolii Karatsuba

Instructor: [Explanation of what trees are]

Instructor: Back when you were studying grammar, how many people built parse trees?
Class: No [but 3-4 hands raised]

Class: <no>
Class: no
Class: no
Class: perhaps
Class: c-c-combo breaker
Class: its super effective!

[Independently of discussion]

Class: to troll or not to troll…
Class: lol wut?
Class: 1. Become Computer Scientist

Class: 2. Build robot army.
I SEE WHAT YOU DID THERE
Metatree: a tree of trees [Tangential relation to topic]
A WILD PIGEON APPEARS!

Instructor: [Discussing removing leaves of a tree]
Also because it is fall

[No reference in conversation]

Instructor: Is a node an ancestor of itself?
Class: [no response]
msosleepy: m so sleepy
[Shortened URL, to Rick Astley Video]
It's a trap
i dont like homeworks :(

Instructor: [Talking about tricky problems on exams and quizzes]
foul trickery is afoot
Instructor: yeah, foul trickery. I don't like problems without already having applied the skill set. I also don't like laundry list type problems
me too

Instructor: [Explanation of tree depth]
wow
What is the height again?
The maximum depth of the tree. You can count the Levels by generation
Not 5? It's max level and not count?
Instructor: Yeah, it's 4 not 5 …[continues on 0 based counting]
In cs you start counting at zero :)

122
Instructor: [Discussing sorted trees]

: What are you referring to with this?

Instructor: [defines binary search trees as a subset of all trees]

: can you draw the picture for the tree after applying rule 3
: :( 
: +1

: How is it useful to have a tree with mul

---

Session 6

Instructor: [Fill out some information online]

(4) : 

: Why do we have an exam the week after halloween?

Instructor: Yeah, there's no good time for midterms

Instructor: It's a great time to do induction... didn't even get a groan, it must be too early in the morning

: ug

: NYEAHHHHHHHHHHHHH GUCCI MANE LA FLAR

: *WHISPER* HEH HEH HE

: how did you get M parents

Instructor: Some of the facts have really easy proofs if you get the trick

: 

: HATE HATE HATE HATE HATE HAT
: I DON'T CARE WHAT THESE CHICKS SA
: I DON'T EVEN LOOK THAT WAY
: EVERY TIME I WALK IN THE CLU
: THEY HATIN ON ME CUSE THEY KNO I UK GOO

Instructor: The height of the whole subtree, is it thru that it's k – 1?

Class: mumble
APPENDIX D

Distorting Feedback:
Discussion Questions

The following questions were used in generate conversation in the Distorting Feedback study. There were two studies that used these questions. In the first study, groups of three discussed a single question for each conversation. In the second study, pairs discussed three questions for each conversation. Some questions were added and one was removed between the two studies. Questions that appeared by themselves in the first study are marked with a †. Questions from the follow up study appear in the set (A, B, C, D) of questions that they appeared with in the study.

Only in the group experiment
† Would you be willing to murder an innocent person if it would end hunger in the world?

Question Set A
† If you could spend one year in perfect happiness but afterward would remember nothing of the experience would you do so? If not, Why not?

Which is more important: actual experiences, or the memories that remain when the experiences are over?
† If you could take a one-month trip anywhere in the world and money were not a consideration, where would you go and what would you do?
If you knew there would be a nuclear war in one week, what would you do?

Question Set B
Do you feel ill at ease going alone to either dinner or a movie? What about going on a vacation by yourself?
† Would you prefer to be blind or deaf? How would it separate you from others?
How would it affect your job and hobbies?
† Given the ability to project yourself into the future but not return, would you do so? If not, would you change your mind if you could take someone along?
How far would you go?
What would induce you to give up life as you know it and face the unknown? Were people in previous centuries more adventurous than we are today or was it simply harder to avoid risk and adventure? How much does affluence make people complacent and averse to risk?

**Question Set C**

If a crystal ball would tell you the truth about any one thing you wished to know concerning yourself, life, the future, or anything else, what would you want to know?

† If you were able to live to the age of ninety and retain either the body or the mind of a thirty year old while the other aged naturally for the last sixty years of your life, which would you choose?

† Would you like to know the precise date of your death? How would this affect your daily life?

*What if everyone knew their date of death?*

**Question Set D**

Would you like to be famous? In what way?

Would you accept $1,000,000 to leave the country and never set foot in it again?

*If you were expelled from the country and had only limited financial resources, where and how would you try to rebuild your life?*

† You are leading 100 people whose lives are in danger and you must choose between two courses of action. One would save only ninety people; the other would have some chance of saving everyone but were it to fail everyone would die. Which would you choose?

*What if you had to choose the ten people who would die? Would you rather have someone else in the group make the decision even though you might be picked to die?*
Distorting Feedback:
Extended Tables

### Table E.1: Collocated groups, contrasting Talkative and Quiet participants

| Measure | $F_{1,24}$ | Sig | Talkative | | | | | Quiet | | | |
|---------|-----------|-----|----------|----|---|---|----|---|---|---|---|---|
| Leads   | 14.4      | $p < 0.001$ | 17.3 | 4.7 | 14.0 | 5.0 | | | | | | |
| Turns   | 0.1       | $p < 0.9$ | 7.8  | 2.0 | 7.2  | 1.8 | | | | | | |
| Length  | 11.8      | $p < 0.002$ | 10.4 | 3.0 | 9.5  | 3.4 | | | | | | |
| Q1      | 2.9       | $p < 0.1$ | 4.0  | 1.3 | 3.8  | 1.4 | | | | | | |
| Q2      | 0.2       | $p < 0.7$ | 3.9  | 0.9 | 3.9  | 1.3 | | | | | | |
| Q3      | 0.2       | $p < 0.7$ | 2.6  | 1.4 | 2.6  | 1.7 | | | | | | |
| Q4      | 3.0       | $p < 0.1$ | 5.5  | 1.6 | 5.9  | 1.4 | | | | | | |
| Q5      | 1.7       | $p < 0.2$ | 5.1  | 1.7 | 5.4  | 1.8 | | | | | | |
| Q6      | 0.0       | $p < 0.9$ | 4.8  | 1.9 | 4.8  | 2.0 | | | | | | |
| Q7      | 0.6       | $p < 0.5$ | 2.3  | 1.4 | 2.1  | 1.2 | | | | | | |
| Q8      | 2.5       | $p < 0.2$ | 2.3  | 1.4 | 1.9  | 1.2 | | | | | | |
| Q9      | 0.1       | $p < 0.9$ | 2.4  | 1.4 | 2.0  | 1.4 | | | | | | |

### Table E.2: Collocated groups, contrasting states of emphasis

| Measure | $F_{2,48}$ | Sig | Normal | | | | | Emphasized | | | | | | Under-Emph | | | | |
|---------|-----------|-----|--------|----|---|---|----|---|---|---|---|---|---|---|---|---|
| Leads   | 0.9       | $p < 0.5$ | 15.5 | 5.4 | 14.7 | 4.8 | 15.7 | 5.3 | | | | | | | | |
| Turns   | 1.6       | $p < 0.2$ | 7.5  | 1.8 | 7.4  | 2.0 | 7.4  | 1.9 | | | | | | | | |
| Length  | 1.2       | $p < 0.4$ | 10.0 | 3.5 | 9.4  | 3.2 | 10.0 | 3.5 | | | | | | | | |
| Q1      | 5.4       | $p < 0.0001$ | 4.0  | 1.4 | 4.3  | 1.3 | 3.5  | 1.3 | | | | | | | | |
| Q2      | 1.0       | $p < 0.4$ | 3.9  | 1.4 | 4.1  | 1.1 | 3.8  | 1.1 | | | | | | | | |
| Q3      | 1.1       | $p < 0.4$ | 2.9  | 1.7 | 2.4  | 1.5 | 2.6  | 1.5 | | | | | | | | |
| Q4      | 1.5       | $p < 0.2$ | 5.5  | 1.6 | 5.8  | 1.4 | 5.7  | 1.5 | | | | | | | | |
| Q5      | 0.9       | $p < 0.5$ | 5.1  | 1.8 | 5.3  | 1.9 | 5.3  | 1.6 | | | | | | | | |
| Q6      | 0.3       | $p < 0.9$ | 4.6  | 2.0 | 4.9  | 2.0 | 4.8  | 2.0 | | | | | | | | |
| Q7      | 0.4       | $p < 0.8$ | 2.0  | 1.1 | 2.2  | 1.2 | 2.2  | 1.4 | | | | | | | | |
| Q8      | 1.0       | $p < 0.4$ | 2.2  | 1.4 | 2.0  | 1.3 | 2.1  | 1.3 | | | | | | | | |
| Q9      | 1.8       | $p < 0.1$ | 2.3  | 1.5 | 2.3  | 1.5 | 2.2  | 1.4 | | | | | | | | |
Table E.3: Collocated groups, contrasting Talkative and Quiet participants in states of emphasis

<table>
<thead>
<tr>
<th>Measure</th>
<th>$F_{2.48}$</th>
<th>Sig</th>
<th>Category</th>
<th>Normal $\mu$</th>
<th>Normal $\sigma$</th>
<th>Emphasized $\mu$</th>
<th>Emphasized $\sigma$</th>
<th>Under-Emph $\mu$</th>
<th>Under-Emph $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leads</td>
<td>2.7</td>
<td>$p &lt; 0.02$</td>
<td>Talkative</td>
<td>18.0</td>
<td>4.9</td>
<td>16.5</td>
<td>4.4</td>
<td>17.1</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>13.6</td>
<td>5.0</td>
<td>13.4</td>
<td>4.6</td>
<td>14.6</td>
<td>5.3</td>
</tr>
<tr>
<td>Turns</td>
<td>1.5</td>
<td>$p &lt; 0.2$</td>
<td>Talkative</td>
<td>7.6</td>
<td>1.8</td>
<td>8.0</td>
<td>2.2</td>
<td>7.6</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>7.4</td>
<td>1.7</td>
<td>7.0</td>
<td>1.8</td>
<td>7.2</td>
<td>1.7</td>
</tr>
<tr>
<td>Length</td>
<td>1.3</td>
<td>$p &lt; 0.3$</td>
<td>Talkative</td>
<td>11.1</td>
<td>3.6</td>
<td>9.6</td>
<td>3.0</td>
<td>10.6</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>6.1</td>
<td>3.31</td>
<td>9.4</td>
<td>3.3</td>
<td>9.7</td>
<td>3.2</td>
</tr>
<tr>
<td>Q1</td>
<td>2.3</td>
<td>$p &lt; 0.04$</td>
<td>Talkative</td>
<td>4.3</td>
<td>1.2</td>
<td>4.1</td>
<td>1.2</td>
<td>3.6</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>3.7</td>
<td>1.4</td>
<td>4.3</td>
<td>1.3</td>
<td>3.4</td>
<td>1.3</td>
</tr>
<tr>
<td>Q2</td>
<td>1.0</td>
<td>$p &lt; 0.5$</td>
<td>Talkative</td>
<td>4.0</td>
<td>1.0</td>
<td>3.9</td>
<td>0.8</td>
<td>3.8</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>3.8</td>
<td>1.6</td>
<td>4.3</td>
<td>1.1</td>
<td>3.7</td>
<td>1.2</td>
</tr>
<tr>
<td>Q3</td>
<td>1.1</td>
<td>$p &lt; 0.4$</td>
<td>Talkative</td>
<td>2.4</td>
<td>1.3</td>
<td>2.3</td>
<td>1.1</td>
<td>2.8</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>3.1</td>
<td>1.9</td>
<td>2.4</td>
<td>1.7</td>
<td>2.4</td>
<td>1.5</td>
</tr>
<tr>
<td>Q4</td>
<td>1.3</td>
<td>$p &lt; 0.3$</td>
<td>Talkative</td>
<td>5.7</td>
<td>1.5</td>
<td>5.5</td>
<td>1.3</td>
<td>5.5</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>5.3</td>
<td>1.6</td>
<td>6.0</td>
<td>1.3</td>
<td>5.8</td>
<td>1.3</td>
</tr>
<tr>
<td>Q5</td>
<td>0.6</td>
<td>$p &lt; 0.7$</td>
<td>Talkative</td>
<td>5.2</td>
<td>1.6</td>
<td>4.9</td>
<td>1.9</td>
<td>5.1</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>5.0</td>
<td>1.8</td>
<td>5.4</td>
<td>1.8</td>
<td>5.5</td>
<td>1.6</td>
</tr>
<tr>
<td>Q6</td>
<td>0.3</td>
<td>$p &lt; 0.9$</td>
<td>Talkative</td>
<td>4.7</td>
<td>2.0</td>
<td>4.9</td>
<td>1.8</td>
<td>4.7</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>4.4</td>
<td>2.0</td>
<td>4.8</td>
<td>2.0</td>
<td>4.8</td>
<td>2.0</td>
</tr>
<tr>
<td>Q7</td>
<td>0.7</td>
<td>$p &lt; 0.6$</td>
<td>Talkative</td>
<td>1.8</td>
<td>1.0</td>
<td>2.4</td>
<td>1.2</td>
<td>2.3</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>2.1</td>
<td>1.1</td>
<td>1.9</td>
<td>1.1</td>
<td>2.0</td>
<td>1.3</td>
</tr>
<tr>
<td>Q8</td>
<td>0.3</td>
<td>$p &lt; 0.9$</td>
<td>Talkative</td>
<td>2.2</td>
<td>1.2</td>
<td>2.2</td>
<td>1.4</td>
<td>2.3</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>2.1</td>
<td>1.4</td>
<td>1.8</td>
<td>1.1</td>
<td>1.8</td>
<td>1.1</td>
</tr>
<tr>
<td>Q9</td>
<td>1.3</td>
<td>$p &lt; 0.5$</td>
<td>Talkative</td>
<td>2.9</td>
<td>1.5</td>
<td>2.3</td>
<td>1.5</td>
<td>2.4</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>1.9</td>
<td>1.3</td>
<td>2.2</td>
<td>1.5</td>
<td>2.0</td>
<td>1.3</td>
</tr>
</tbody>
</table>

Table E.4: Remote pairs, contrasting Talkative and Quiet participants

<table>
<thead>
<tr>
<th>Measure</th>
<th>$F_{1.24}$</th>
<th>Sig</th>
<th>Talkative $\mu$</th>
<th>Talkative $\sigma$</th>
<th>Quiet $\mu$</th>
<th>Quiet $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leads</td>
<td>9.8</td>
<td>$p &lt; 0.005$</td>
<td>26.7</td>
<td>4.3</td>
<td>22.1</td>
<td>4.3</td>
</tr>
<tr>
<td>Turns</td>
<td>0.6</td>
<td>$p &lt; 0.5$</td>
<td>6.4</td>
<td>1.8</td>
<td>6.9</td>
<td>1.8</td>
</tr>
<tr>
<td>Length</td>
<td>6.4</td>
<td>$p &lt; 0.02$</td>
<td>20.7</td>
<td>7.7</td>
<td>15.4</td>
<td>4.4</td>
</tr>
<tr>
<td>Q1'</td>
<td>6.2</td>
<td>$p &lt; 0.02$</td>
<td>4.4</td>
<td>1.2</td>
<td>3.7</td>
<td>1.1</td>
</tr>
<tr>
<td>Q2'</td>
<td>1.2</td>
<td>$p &lt; 0.3$</td>
<td>4.3</td>
<td>1.2</td>
<td>4.0</td>
<td>0.9</td>
</tr>
<tr>
<td>Q3'</td>
<td>0.2</td>
<td>$p &lt; 0.7$</td>
<td>3.6</td>
<td>1.1</td>
<td>3.4</td>
<td>1.2</td>
</tr>
<tr>
<td>Q4'</td>
<td>1.0</td>
<td>$p &lt; 0.4$</td>
<td>3.3</td>
<td>1.1</td>
<td>3.8</td>
<td>1.6</td>
</tr>
<tr>
<td>Q5'</td>
<td>2.0</td>
<td>$p &lt; 0.2$</td>
<td>3.7</td>
<td>1.0</td>
<td>3.2</td>
<td>1.4</td>
</tr>
<tr>
<td>Q6'</td>
<td>0.5</td>
<td>$p &lt; 0.5$</td>
<td>5.1</td>
<td>1.4</td>
<td>5.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Q7'</td>
<td>0.1</td>
<td>$p &lt; 0.7$</td>
<td>5.4</td>
<td>1.4</td>
<td>5.5</td>
<td>1.6</td>
</tr>
<tr>
<td>Q8'</td>
<td>0.0</td>
<td>$p &lt; 0.9$</td>
<td>4.9</td>
<td>1.6</td>
<td>5.0</td>
<td>1.8</td>
</tr>
<tr>
<td>Q9'</td>
<td>0.5</td>
<td>$p &lt; 0.5$</td>
<td>4.9</td>
<td>1.6</td>
<td>4.5</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Table E.4: Remote pairs, contrasting Talkative and Quiet participants
Table E.5: Remote pairs, contrasting states of emphasis

<table>
<thead>
<tr>
<th>Measure</th>
<th>$F_{2,48}$</th>
<th>Sig</th>
<th>Normal</th>
<th>Emphasized</th>
<th>Under-Emph</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\mu$</td>
<td>$\sigma$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>Leads</td>
<td>4.2</td>
<td>$p &lt; 0.02$</td>
<td>24.6</td>
<td>4.9</td>
<td>23.2</td>
</tr>
<tr>
<td>Turns</td>
<td>16.2</td>
<td>$p &lt; 0.0001$</td>
<td>5.8</td>
<td>1.4</td>
<td>7.3</td>
</tr>
<tr>
<td>Length</td>
<td>1.3</td>
<td>$p &lt; 0.3$</td>
<td>17.9</td>
<td>6.2</td>
<td>19.2</td>
</tr>
<tr>
<td>$Q_1'$</td>
<td>0.8</td>
<td>$p &lt; 0.5$</td>
<td>4.2</td>
<td>1.2</td>
<td>4.0</td>
</tr>
<tr>
<td>$Q_2'$</td>
<td>1.6</td>
<td>$p &lt; 0.3$</td>
<td>4.4</td>
<td>0.9</td>
<td>4.1</td>
</tr>
<tr>
<td>$Q_3'$</td>
<td>0.5</td>
<td>$p &lt; 0.7$</td>
<td>3.6</td>
<td>1.1</td>
<td>3.3</td>
</tr>
<tr>
<td>$Q_4'$</td>
<td>0.4</td>
<td>$p &lt; 0.7$</td>
<td>3.5</td>
<td>1.3</td>
<td>3.5</td>
</tr>
<tr>
<td>$Q_5'$</td>
<td>0.6</td>
<td>$p &lt; 0.6$</td>
<td>3.2</td>
<td>1.2</td>
<td>3.5</td>
</tr>
<tr>
<td>$Q_6'$</td>
<td>0.4</td>
<td>$p &lt; 0.7$</td>
<td>5.2</td>
<td>1.4</td>
<td>5.3</td>
</tr>
<tr>
<td>$Q_7'$</td>
<td>0.0</td>
<td>$p &lt; 1$</td>
<td>5.5</td>
<td>1.4</td>
<td>5.5</td>
</tr>
<tr>
<td>$Q_8'$</td>
<td>0.6</td>
<td>$p &lt; 0.6$</td>
<td>4.7</td>
<td>1.7</td>
<td>5.1</td>
</tr>
<tr>
<td>$Q_9'$</td>
<td>0.1</td>
<td>$p &lt; 0.9$</td>
<td>4.6</td>
<td>1.8</td>
<td>4.7</td>
</tr>
</tbody>
</table>

Table E.6: Remote pairs, contrasting Talkative and Quiet participants in states of emphasis

<table>
<thead>
<tr>
<th>Measure</th>
<th>$F_{2,48}$</th>
<th>Sig</th>
<th>Category</th>
<th>Normal</th>
<th>Emphasized</th>
<th>Under-Emph</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\mu$</td>
<td>$\sigma$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>Leads</td>
<td>0.5</td>
<td>$p &lt; 0.6$</td>
<td>Talkative</td>
<td>27.3</td>
<td>3.7</td>
<td>25.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>21.9</td>
<td>4.5</td>
<td>21.0</td>
</tr>
<tr>
<td>Turns</td>
<td>0.7</td>
<td>$p &lt; 0.6$</td>
<td>Talkative</td>
<td>5.6</td>
<td>1.2</td>
<td>6.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>6.0</td>
<td>1.6</td>
<td>7.7</td>
</tr>
<tr>
<td>Length</td>
<td>0.5</td>
<td>$p &lt; 0.6$</td>
<td>Talkative</td>
<td>20.5</td>
<td>6.7</td>
<td>22.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>15.4</td>
<td>4.6</td>
<td>15.8</td>
</tr>
<tr>
<td>$Q_1'$</td>
<td>0.5</td>
<td>$p &lt; 0.6$</td>
<td>Talkative</td>
<td>4.7</td>
<td>1.2</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>3.7</td>
<td>1.1</td>
<td>3.9</td>
</tr>
<tr>
<td>$Q_2'$</td>
<td>0.2</td>
<td>$p &lt; 0.8$</td>
<td>Talkative</td>
<td>4.6</td>
<td>1.1</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>4.2</td>
<td>0.8</td>
<td>3.9</td>
</tr>
<tr>
<td>$Q_3'$</td>
<td>0.2</td>
<td>$p &lt; 0.9$</td>
<td>Talkative</td>
<td>3.7</td>
<td>1.1</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>3.6</td>
<td>1.1</td>
<td>3.2</td>
</tr>
<tr>
<td>$Q_4'$</td>
<td>0.4</td>
<td>$p &lt; 0.7$</td>
<td>Talkative</td>
<td>3.3</td>
<td>0.8</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>3.6</td>
<td>1.8</td>
<td>3.6</td>
</tr>
<tr>
<td>$Q_5'$</td>
<td>0.8</td>
<td>$p &lt; 0.5$</td>
<td>Talkative</td>
<td>3.5</td>
<td>0.9</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>3.0</td>
<td>1.4</td>
<td>3.1</td>
</tr>
<tr>
<td>$Q_6'$</td>
<td>0.4</td>
<td>$p &lt; 0.7$</td>
<td>Talkative</td>
<td>5.0</td>
<td>0.9</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>5.4</td>
<td>1.8</td>
<td>5.3</td>
</tr>
<tr>
<td>$Q_7'$</td>
<td>0.3</td>
<td>$p &lt; 0.8$</td>
<td>Talkative</td>
<td>5.4</td>
<td>0.9</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>5.5</td>
<td>1.8</td>
<td>5.5</td>
</tr>
<tr>
<td>$Q_8'$</td>
<td>0.5</td>
<td>$p &lt; 0.7$</td>
<td>Talkative</td>
<td>4.9</td>
<td>1.3</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>4.6</td>
<td>2.1</td>
<td>5.2</td>
</tr>
<tr>
<td>$Q_9'$</td>
<td>1.6</td>
<td>$p &lt; 0.3$</td>
<td>Talkative</td>
<td>5.0</td>
<td>1.7</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quiet</td>
<td>4.2</td>
<td>1.8</td>
<td>4.3</td>
</tr>
</tbody>
</table>
References


