

# Hacking Innovation - Group Dynamics in Innovation Teams

by

Oren Lederman

Submitted to the Program in Media Arts and Sciences,  
School of Architecture and Planning,  
in partial fulfillment of the requirements for the degree of  
Master of Science in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2015

© Massachusetts Institute of Technology 2015. All rights reserved.

Author .....  
Program in Media Arts and Sciences  
Aug 06, 2015

Certified by.....  
Alex "Sandy" P. Pentland  
Toshiba Professor of Media Arts and Sciences  
Thesis Supervisor

Accepted by.....  
Pattie Maes  
Academic Head  
Program in Media Arts and Sciences



# Hacking Innovation - Group Dynamics in Innovation Teams

by

Oren Lederman

Submitted to the Program in Media Arts and Sciences,  
School of Architecture and Planning,  
on Aug 06, 2015, in partial fulfillment of the  
requirements for the degree of  
Master of Science in Media Arts and Sciences

## Abstract

Innovative work is gradually shifting towards collaborative efforts by larger, multi-disciplinary teams, making team efficacy an increasingly important field of study. Researchers in this field have mainly focused on laboratory experiments, which may not fully capture the complex situations that teams encounter in real life. The alternative, field studies, are difficult to maintain and often require significant time to produce results.

In this thesis we propose a methodology that bridges the gap between these two settings—the laboratory and the field. By combining a new, affordable electronic badge that captures vocalization data with an innovative setting—the Hackathon—we create a new environment for studying team performance. This methodology reduces the duration and maintenance burden of such studies, and offers new opportunities for examining the effects of interventions on teamwork.

The preliminary results from our studies show a variety of individual and team behaviors that can be captured in Hackathons using badges, such as participation, the parity of contribution to group discussions, the level of turn taking, and the frequency and duration of meetings. In a Hackathon, we measure these behaviors throughout the entire life cycle of each team, observe how they change in response to different shocks, and study how well the team members collaborate and perform as a team.

Thesis Supervisor: Alex "Sandy" P. Pentland  
Title: Toshiba Professor of Media Arts and Sciences



Hacking Innovation - Group Dynamics in Innovation Teams  
by  
Oren Lederman

Thesis Advisor.....  
Alex "Sandy" P. Pentland  
Toshiba Professor of Media Arts and Sciences  
MIT Program in Media Arts and Sciences

Thesis reader.....  
Kent Larson  
Principal Research Scientist  
MIT Media Lab

Thesis reader.....  
Fiona E. Murray  
Professor of Entrepreneurship  
and co-director of MIT Innovation Initiative  
MIT Sloan School of Management



## Acknowledgments

I would first like to thank my advisor, Alex "Sandy" Pentland, for his support and guidance over the last two years. I would like to thank my co-advisor and thesis reader, Fiona Murray, for this opportunity and for her counsel. I am also grateful to Kent Larson, my thesis reader, for his insightful comments and feedback.

I am thankful to Moshe Ben-Akiva, who helped make this happen; to the members of ITS Lab and Future Mobility; and especially to Francisco C. Pereira, for his help and guidance before and throughout the program.

This work would not have been possible without the support of my friends and peers from the Human Dynamics group and the MIT Media Lab. Also, many thanks for my UROPs, Shirley, Devin, and Angus, for their important contributions to this work and my other projects.

Special thanks to my friends and colleagues from the LABi team, in particular Patrick Schneider-Sikorsky for making the introduction, and to Daniel Fehder—buddy, I can never thank you enough for all your help.

Financial support from the Ewing Marion Kauffman Foundation is gratefully acknowledged. I would also like to acknowledge the MIT Media Lab member companies for their financial support throughout my studies.

My final acknowledgment goes to my family for their ongoing support—my parents Gil and Bella, my brother Roy, and my fiancée Deana, for her love, care and support. Brace yourself—we are not done here just yet.





# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>19</b> |
| <b>2</b> | <b>Background</b>   | <b>21</b> |
| 2.1      | Traditional Environments for Team Performance Studies . . . . . | 21        |
| 2.2      | Hackathons as a Potential Research Setting . . . . .            | 22        |
| 2.3      | Electronic Badges . . . . .                                     | 24        |
| 2.3.1    | Technical Overview and Previous Work . . . . .                  | 24        |
| 2.3.2    | Benefits for Research . . . . .                                 | 25        |
| 2.3.3    | Conversational Patterns . . . . .                               | 27        |
| 2.4      | Team Performance . . . . .                                      | 28        |
| 2.4.1    | Team Composition . . . . .                                      | 28        |
| 2.4.2    | Coordination and Communication . . . . .                        | 29        |
| 2.4.3    | Shared Knowledge . . . . .                                      | 30        |
| 2.4.4    | Matching Expectations . . . . .                                 | 31        |
| <b>3</b> | <b>Hardware Design</b>  | <b>33</b> |
| 3.1      | Initial Requirements . . . . .                                  | 33        |
| 3.2      | Functional Description . . . . .                                | 34        |
| 3.3      | Badge Design . . . . .  | 35        |
| 3.3.1    | Bluetooth Low-Energy . . . . .                                  | 36        |
| 3.3.2    | Microphone . . . . .  | 38        |
| 3.3.3    | Microphone Low-pass Filter . . . . .                            | 38        |
| 3.3.4    | Power Supply . . . . .  | 39        |

|          |  |           |
|----------|--|-----------|
| 3.3.5    | Microphone Voltage Regulator . . . . .                                       | 39        |
| 3.4      | Base Station . . . . .   | 39        |
| <b>4</b> | <b>Data Analysis</b>   | <b>43</b> |
| 4.1      | Determining Badge Volume . . . . .   | 43        |
| 4.2      | Detecting the Active Speaker . . . . .                                       | 44        |
| 4.3      | Computing Turn Taking . . . . .  | 44        |
| <b>5</b> | <b>Badge Evaluation</b>  | <b>47</b> |
| 5.1      | The Marshmallow Challenge . . . . .  | 48        |
| 5.1.1    | Background . . . . .   | 48        |
| 5.1.2    | Study Description . . . . .  | 50        |
| 5.1.3    | Results . . . . .  | 51        |
| 5.2      | Hackathons . . . . .   | 55        |
| 5.2.1    | Background . . . . .   | 55        |
| 5.2.2    | MIT Hacking Medicine—Grand Hack 2015 . . . . .                               | 56        |
| 5.2.3    | Corporate Hackathon By a Multinational Pharmaceutical Com-<br>pany . . . . . | 61        |
| <b>6</b> | <b>Discussion</b>  | <b>77</b> |
| 6.1      | Badge Evaluation . . . . .   | 77        |
| 6.1.1    | Conversation Patterns . . . . .  | 77        |
| 6.1.2    | Limitations . . . . .  | 78        |
| 6.2      | Hackathons as a Research Setting . . . . .                                   | 78        |
| 6.2.1    | Observed Team and Teamwork Characteristics . . . . .                         | 79        |
| 6.2.2    | Potential Interventions . . . . .  | 82        |
| 6.2.3    | Recommendations for Future Studies in Hackathons . . . . .                   | 83        |
| <b>7</b> | <b>Conclusions</b>   | <b>85</b> |
| 7.1      | Future Work . . . . .  | 86        |
| <b>A</b> | <b>Schematic Diagrams and PCB Layout</b>                                     | <b>89</b> |





# List of Figures

|     |  |    |
|-----|--|----|
| 2-1 | Sociometric badge evolution . . . . .  | 26 |
| 3-1 | A simple deployment that includes two badges, a base-station, and a portable Wi-Fi access point . . . . .                    | 35 |
| 3-2 | A badge carried inside a convention badge holder . . . . .   | 37 |
| 3-3 | A badge with a neck cord . . . . .   | 37 |
| 3-4 | A Raspberry Pi base station . . . . .  | 41 |
| 4-1 | Raw vocalization samples from two badges. Badge #4 is reading a monologue between 31:15 and 32:08 . . . . .                  | 45 |
| 4-2 | Monologue data after MA . . . . .  | 45 |
| 4-3 | Speaker detection. Red shows badge #2 as the speaker. Blue shows badge #4 as the speaker. Black represents silence . . . . . | 45 |
| 5-1 | Performance of different types of teams. Source: Tom Wujec, TED April 2010 . . . . .   | 49 |
| 5-2 | Team #1, percentage of conversation time per minute . . . . .  | 52 |
| 5-3 | Team #1, number of turns taken per minute . . . . .  | 52 |
| 5-4 | Team #2, percentage of conversation time per minute . . . . .  | 53 |
| 5-5 | Team #2, number of turns taken per minute . . . . .  | 53 |
| 5-6 | Team #1, Contribution of each badge to the total conversation time (%)   | 54 |
| 5-7 | Team #2, Contribution of each badge to the total conversation time (%)   | 54 |
| 5-8 | Team #1, Percentage of conversation time per minute, day 1 . . . . .   | 59 |
| 5-9 | Team #1, Number of turns taken per minute, day 1 . . . . .   | 59 |

|      |   |    |
|------|---|----|
| 5-10 | Team #1, Percentage of conversation time per minute, day 2 . . . . .  | 59 |
| 5-11 | Team #1, Number of turns taken per minute, day 2 . . . . .  | 60 |
| 5-12 | Team #1, Contribution of each badge to the total conversation time<br>(%), day 1 . . . . .  | 60 |
| 5-13 | Team #1, Contribution of each badge to the total conversation time<br>(%), day 2 . . . . .  | 60 |
| 5-14 | Team #2, number of samples collected by each badge during the first<br>day . . . . .  | 61 |
| 5-15 | Team #2, number of samples collected by each badge during the second<br>day . . . . .   | 62 |
| 5-16 | Percentage of conversation time per minute. Corporate Hackathon,<br>Day 1, 9:30 am - 1:00 pm . . . . .                                  | 64 |
| 5-17 | Number of turns taken per minute. Corporate Hackathon, Day 1,<br>9:30 am - 1:00 pm . . . . .  | 65 |
| 5-18 | Contribution (in seconds) to the conversation, by participant. Corpo-<br>rate Hackathon, Day 1, 9:30 am - 1:00 pm . . . . .             | 65 |
| 5-19 | Percentage of conversation time per minute. Corporate Hackathon,<br>Day 1, 11:15 am - 12:00 pm . . . . .                                | 68 |
| 5-20 | Number of turns taken per minute. Corporate Hackathon, Day 1,<br>11:15 am - 12:00 pm . . . . .  | 68 |
| 5-21 | Group #1, Contribution (in seconds) to the conversation, by partici-<br>pant. Corporate Hackathon, Day 1, 11:15 am - 12:00 pm . . . . . | 69 |
| 5-22 | Group #2, Contribution (in seconds) to the conversation, by partici-<br>pant. Corporate Hackathon, Day 1, 11:15 am - 12:00 pm . . . . . | 69 |
| 5-23 | Percentage of conversation time per minute. Corporate Hackathon,<br>Day 1, 1:00 pm - 5:30 pm . . . . .                                  | 71 |
| 5-24 | Number of turns taken per minute. Corporate Hackathon, Day 1,<br>1:00 pm - 5:30 pm . . . . .  | 71 |
| 5-25 | Contribution (in seconds) to the conversation, by participant. Corpo-<br>rate Hackathon, Day 1, 1:00 pm - 5:30 pm . . . . .             | 72 |

|      |   |    |
|------|---|----|
| 5-26 | Percentage of conversation time per minute. Corporate Hackathon,<br>Day 2, 9:20 am - 4:45 pm . . . . .                      | 74 |
| 5-27 | Number of turns taken per minute. Corporate Hackathon, Day 2,<br>9:20 am - 4:45 pm . . . . .                                | 74 |
| 5-28 | Contribution (in seconds) to the conversation, by participant. Corpo-<br>rate Hackathon, Day 2, 9:20 am - 4:45 pm . . . . . | 75 |
| A-1  | Badge PCB schematic . . . . .   | 90 |
| A-2  | Badge PCB top layer layout . . . . .  | 91 |
| A-3  | Badge PCB bottom layer layout . . . . .   | 92 |
| B-1  | Bill of Materials . . . . .   | 94 |





# List of Tables

|                                   |    |
|-----------------------------------|----|
| 3.1 Popular BLE modules . . . . . | 37 |
|-----------------------------------|----|



# Chapter 1

## Introduction

Innovation has gradually shifted from the work of lone geniuses towards collaboration; large, multi-disciplinary teams make more important scientific discoveries than 30 years ago and teams are increasingly the prevailing work unit in corporations [90]. This trend has led to a growing interest in studying the factors affecting team efficacy. Researchers have been trying to answer questions such as: What group dynamics lead to better performance? How does the composition of a team affect collaboration, access to information and the team's ability to generate innovative ideas? How do interventions lead to better performance?

To answer these questions, researchers have conducted experiments to measure the effects of different factors on team performance in well-defined tasks, mostly in the laboratory setting [89, 56]. However, as some of this research suggests, teams do not necessarily follow a well-structured, systematic process [44, 35, 36], and laboratory experiments do not fully capture the complex situations that teams encounter in real life [66]. The alternative, field studies, are difficult to maintain and often require significant time to produce results.

Wearable devices, such as the Sociometric Badges developed in the Human Dynamics group at the MIT Media Lab [65], help reduce the burden and bias involved in such experiments by automatically quantifying social behavior [70, 89]. Reducing the cost of this technology and making it more available would allow researchers to deploy it on a large scale.

In this thesis we propose a new methodology that simplifies and shortens the time it takes to conduct team performance studies. This methodology uses a new, affordable wearable device to collect data on the behavior of teams in Hackathons—innovative settings that combine a short lifespan with the characteristics of a real working environment. An entry survey with questions regarding demographics and personal background and the Ten-Item Personality Inventory (TIPI) [37] of the Big-Five personality dimensions will provide us with information about the composition of the teams.

To evaluate the new badges, we first tested them in a short team-building exercise called "The Marshmallow Challenge"—an activity in which participants are asked to work together to build the tallest freestanding tower made of spaghetti sticks and tape that can support a marshmallow at the top. In Section 5.1 we describe these tests and explain the similarities between teamwork in this exercise and in Hackathons.

We then conducted two Hackathon studies to test the stability of the data collection platform and acceptance by participants, and to evaluate Hackathons as a research setting. Section 5.2 describes these studies.

Due to time and cost limitations we have not yet made a large scale deployment of the system, and instead evaluated the Hackathon as a setting for our future studies.

We begin with Chapter 2—an overview of the background and previous work related to the research topic. Next, we provide detailed information on the data collection platform and the considerations that guided its design in Chapter 3. In Chapter 4 we describe the methods used to transform raw data into actionable products. In Chapter 5 we present the studies we used to evaluate the performance of badges and the potential of Hackathons as a research setting. In Chapter 6 we discuss the results of these evaluations, and finally in Chapter 7 we present the conclusions and plans for future work.

# Chapter 2

## Background

### 2.1 Traditional Environments for Team Performance Studies

Traditional studies mostly focus on laboratory experiments and individual, well-defined tasks. For example, in [89], groups were asked to perform different well-defined tasks one at a time to measure their performance; in one task they were given a brick and asked to come up with as many uses for it as possible, to see how well they brainstorm as a group, and on the next task they were asked to simultaneously type as much text as they could in a shared document to see how well they execute a task together. In a field experiment, Lim & Klein studied the effect on the performance of combat units in various well-defined tasks, such as securing a bridge or laying an ambush [56].

Because of the complexity of real-world working processes, some researchers claim that these types of experiments do not represent the real working environment of the team. A number of studies suggest that teams neither follow a universal series of stages, nor benefit from a well-structured, systematic approach to problem solving [44, 35, 36]. Instead, teams choose a strategy and stick to it for half of their time together, then switch to a different strategy. The transition between the two phases is effected by the synchrony in group members' expectations about deadlines. This

behavior was observed both in naturally-occurring teams [35] and in the lab [36].

These results suggest that the factors affecting team efficacy should be tested and measured not just for one task at a time, but for the entire process. Moreover, teams are expected to go through the full range of tasks, constantly making decisions on how to go about these tasks, and not just on the task itself. Therefore, team performance might be affected by how much the team members agree with and understand the current status, how much they agree on what they should be doing next, and how well they coordinate their work.

O'Reilly, Williams and Barsade [66] suggested field studies as an alternative: “as a topic, diversity in organizations and its consequences is too important to relegate only to the laboratory. Even though laboratory studies allow us to vary the types of diversity we are interested in, while holding constant other variables, they do not fully capture the complex situations diverse groups are faced with in organizations. In the field, we are able to capture the characteristics of work groups that are impossible to maintain in the laboratory setting” . We extend this claim to the work of innovation teams in general due to the uncertain nature of their work and ambiguous task definitions.

However, field studies face several other challenges—observing a large number of teams requires considerable resources and it takes a long period of time to collect a significant data set.

## 2.2 Hackathons as a Potential Research Setting

Hackathons hold the potential to bridge the gap between the laboratory setting and the field, as they combine short lifespans with the characteristics of the real working environments of innovation teams.

Hackathons are fast-paced events organized by non-profit organizations, universities, corporations, and online communities. They invite participants from different demographics and backgrounds to collaborate, generate new ideas and solve problems. Most Hackathons are open to the public, but others are private, open only

to the employees of the company that organizes them [30, 20, 53, 61, 72, 54]. The outcome of their work can be quantified based on scores provided by the judges, or, in a corporate Hackathon, how many resources are eventually allocated to the ideas.

Traditionally, Hackathons are 24-48 hour events. A presentation covering the subjects and goals of the Hackathon opens the event, followed by a pitch stage in which participants can describe their ideas in one minute. Once the pitch stage is over, participants are free to discuss these ideas and form teams. They then spend most of the remaining time on raw implementation of their ideas, either as working prototypes or as business plans. Finally, each group pitches its idea to the judges, who assess them and announce the winning teams.

Hackathons share several characteristics with real-world innovation environments, such as start-ups and R&D departments:

- In the short bursts of creative work observed in a Hackathon, people are expected to go through the full range of tasks and phases that make up the innovation process[35, 36]
- Teams are composed of individuals who have high degrees of expertise in particular areas, requiring that information contributed by different team members converge in support of critical decisions[41, 42, 82]
- A decision environment with severe time pressure; complex, multicomponent decision tasks; rapidly evolving and changing information; high short-term memory demands and high information ambiguity[83, 51, 52, 63]
- Much like company employees, Hackathon participants are motivated to contribute to the team effort since the outcome of the project may directly affect their futures—it ensures their employment, creates new opportunities, and improves their professional reputations

These characteristics allow us to explore the nature of teamwork when it is separated from some of the organizational forces that shape project selection and team

formation in firms. They allow us to ask some key questions: How do groups come together to work on new, innovative ideas? Do behavioral traits that facilitate forming a group, such as extraversion, help or hinder group dynamics? How can we ensure that teams can tap into the full set of skills, knowledge and creativity of all team members? How can we structure these events to match the specific needs of an organization?

Although Hackathon participants are self-selected, focusing on corporate Hackathons partially solves this issue since all of the participants in these events are already company employees.

## 2.3 Electronic Badges

Wearable devices are becoming a common research tool for studying social behavior. In this section, we present some of the previous studies that experimented with wearable devices and explain the benefits of using them for social studies.

### 2.3.1 Technical Overview and Previous Work

The *iBadge* [67], one of the earliest platforms, used a suite of sensors to collect data and explore how kindergartners interact with each other and with their environments. Each wearable unit weighed 65 grams and had an average battery life of 4.5 hours. It was equipped with a microphone, a DSP unit, a location tracker, a tilt/orientation unit, a temperature sensor, a humidity sensor, a light sensor and a pressure sensor.

In a more recent example, the developers of the OpenBeacon project [8] created a wearable unit called Sensor Tag to gather data on face-to-face interaction patterns between people in different locations, such as workshops [17], conferences [79], hospitals [84] and schools [33]. The latest version of the Sensor Tags uses a PIC16F688 microcontroller and the Nordic nRF51822 Bluetooth Low Energy module to make the badges aware of each other, assess their proximity to each other, and transmit the data back to a base station.

Our research group has been developing wearable devices and mobile frameworks that measure and record face-to-face interaction, location and non-linguistic social



signals [68, 71]. Starting with our early work on the SocioMeter [24] through the more recent Sociometric Badges and Social fMRI frameworks [13, 64, 65], we have explored the use of sensory data and non-linguistic social signals in order to study human behavior and social interaction.

The SocioMeter, a wearable device designed to measure face-to-face interactions, included an IR transceiver, accelerometers and a microphone [24]. A shoulder mount placed the SocioMeter several inches below the wearer’s mouth and reduced the noise created by clothing and movement. With this device, we started developing methods for learning the structure and dynamics of human communication networks.

The Sociometric Badge [64, 65] is the most recent research platform developed by the Human Dynamics group. The second major version uses an Amtel ARM926EJ-S microcontroller capable of running Linux and enough peripherals for all required sensors, with low power consumption. A 2600 mAh battery extends its operational time to 24 hours and the removable microSD HC memory card is capable of storing days of raw data recorded by its digital 3-axis accelerometer, IrDA transceiver and digital MEMS microphone. This system evolved into a commercial solution provided by Sociometric Solutions<sup>®</sup> [11]. Figure 2-1 shows the evolution of the sociometric badges.

### **2.3.2 Benefits for Research**

Wearable devices use unobtrusive sensors to automatically capture the behavior of many individuals, potentially thousands or more [65, 57], at the same time. Compared to direct observations and surveys, this approach produces an objective dataset, and requires fewer human resources, which are expensive.

In our work, smart badges were used for collecting social behavior data such as face-to-face interactions and speaking patterns. The data was then aggregated at the group level to compare group performance in organizations and evaluate design interventions to improve it [65, 70].

In Hackathons, where dozens of teams or more compete at the same time, badges reduce the observational burden by automating the data collection process. This



(a) Sociometer



(b) Sociometric Badge 2.0



(c) Sociometric Badge 3.0

Figure 2-1: Sociometric badge evolution

thesis shows the potential of using Hackathons for studying teams by demonstrating the variability of team behavioral characteristics using the rich vocalization data captured by our wearable devices during these events.

Instrumenting a large number of people with such devices can be costly, especially if we work under the assumption that some participants will not return them. We therefore designed a small and affordable badge that builds upon the extensive experience our research group has in wearable design.

The key for reducing the cost of such a device is to simplify it by keeping only the most valuable sensors and reducing the number of components involved in its manufacturing process. Given the results from our prior studies, we decided to focus on conversational patterns.

### **2.3.3 Conversational Patterns**

Pentland [70, 69] found the patterns of communication to be a strong predictor for team performance in the corporate setting. For example, in a study conducted in a bank call center, Sociometric Badges were used to gather data on the levels of energy (how team members contribute to a team), engagement (how team members communicate with each other) and exploration (how teams communicate with each other) of each team. Analyzing the results, the researchers found that energy and engagement outside of formal meetings explained one-third of the variation in productivity. These results encouraged the center’s managers to change the coffee break schedules so that everyone on the team can take a break at the same time, leading to higher efficiency and higher satisfaction among employees.

Similarly, Woolley et al. used Sociometric Badges to study the Collective Intelligence of groups by instrumenting hundreds of small groups and asking them to perform a series of tasks [89]. These experiments provided evidence for the existence of a general factor explaining groups’ efficacy on a wide variety of tasks. In particular, they found that the parity of vocalization across members of the group is a good proxy for Collective Intelligence.

## 2.4 Team Performance

A large body of research has identified several factors contributing to team efficacy. Here, we review some of the factors related to Hackathon teams, and explain how the data collected with the badges can help us understand these relationships.

### 2.4.1 Team Composition

The background and demographics of the individuals composing the team and their within-group social ties affect the way they collaborate and how well they work as a team. A diverse team may be able to generate novel ideas but the differences can create conflicts and hinder its ability to implement them. Friendship, a history of working together and a high level of social sensitivity have a positive effect on the behavior of the group.

Diversity is a double-edged sword when it comes to team performance. In section 2.4.3 we show that diversity improves performance by giving team members access to rich information and opinions. However, some studies have shown evidence that diversity also increases conflict, reduces cohesion, hinders coordination and effective decision making, and interferes with intra-group communication, all of which have negative effects on team performance [14, 66]. Homogeneity of personality and background were found to have a positive effect on performance, especially in newly formed teams [45, 85].

Familiarity also has positive effects on a team's ability to work together. A group of acquaintances with shared experiences already hold personal knowledge of each other's personalities and skills [73]. These shared experiences help build a common language and social norms, and enable team members to access relevant information more easily [63, 39]. Individuals can also rely on their friends when asked to complete complex tasks [28].

Gender diversity and social sensitivity were found to have a positive relationship with group efficacy. In a study that explored group-level efficacy on a broad range of tasks, researchers showed that there is a group level of intelligence that is uncorrelated

with the individual intelligence of the members of the group. They also identified a key component leading to high correlation across tasks—Collective Intelligence [89]. This component is strongly correlated with psychological and behavioral markers, such as the parity of vocalization across members of the team and the proportion of females in the team.

In Hackathons we expect to see both homogeneous and heterogeneous teams. By combining data from entry surveys and badges, we can study the relationship between team composition and measures of team behavior and success.

## 2.4.2 Coordination and Communication

During a project, the team typically breaks into smaller working units, each focused on a different aspect of the problem. Continuous coordination is essential to reducing uncertainties about others' work progress, establishing shared understanding of the tasks, eliminating redundant work, and ensuring that all of the pieces of the puzzle fit together at the end [26, 21]. Studies have shown that a higher number and greater frequency of interactions facilitates better coordination and positively correlates with high performance [86, 69].

Various collaboration tools, such as shared code repositories, e-mail, and conference calls make it easier for participants to communicate and share their work, but these tools are not as powerful as face-to-face meetings [69, 47, 49]. Face-to-face interaction includes subtle signals that are lost in other means of communication, a source of information crucial for determining others' responsiveness and the acceptability of their decisions [75, 71, 19].

As we show in the next chapters, the creative process of teams in Hackathons typically includes a long opening discussion, followed by a sequence of status meetings. The data collected by our badges allow us to measure the frequency, the duration, and the conversational dynamics of these meetings.

### 2.4.3 Shared Knowledge

In an environment that thrives on innovation, the team’s collective knowledge may determine its ability to solve problems and generate novel ideas. With the increasing pace of knowledge accumulation requiring individuals to spend more time in reaching the knowledge frontier in a field [46], teams enable experts with disparate knowledge to solve complex problems by recombining their unique skills and knowledge in novel ways [82].

The entrepreneurship literature describes the lone entrepreneur who is able to identify an opportunity due to an advantage over the rest of a population—an advantage that emerges from a difference between her own experience and knowledge, and that of others [76].

For a team, this potential arises from the combined knowledge of the team members. A cross-functional team benefits from several advantages over a homogeneous one [14, 66, 87]. First, a team with diversified skills and expertise is expected to have a larger pool of knowledge to tap into, and it will therefore consider a wider range of possible ideas. Second, this diversity produces tension and conflict which contribute to a more complete analysis of the issue at hand, and therefore better decisions.

In contrast, social ties may indicate similar knowledge. In a team with existing ties among its team members, knowledge and opinions are likely to be more redundant than diversified [38], since individuals are more likely to communicate and form friendships with people who share a common language and opinions [39, 15].

The team’s ability to efficiently use the cumulative knowledge of its members is an important factor in turning this potential into a new idea—a successful team can efficiently tap into the combined knowledge of the team members and make use of the unique pieces of information that give the team its advantage. However, this ability may be hindered by different group behaviors, such as production blocking and evaluation apprehension [31].

Production blocking refers to the turn-taking nature of conversations, which prevents people from presenting their ideas as they arise. By the time it is their turn

to speak, the participants have already forgotten some ideas or suppressed them, believing them to be less relevant because a consensus has apparently emerged.

Evaluation apprehension asserts that people might hold back ideas because they fear being evaluated by their teammates, and that the strength of this effect may depend on how they perceive the level of their teammates' expertise.

Teams' tendency to seek similarity may also deter team members from speaking out against the majority opinion and push the team to put more emphasis on common knowledge instead of unique information [87].

Conversation patterns recorded by our badges showed the turn-taking behavior and recorded whether all participants had an opportunity to share information and take an active part in the discussion. An unbalanced contribution in team meetings might indicate inefficient behavior.

#### 2.4.4 Matching Expectations

The degree to which team members understand and agree on their roles and the task plays a part in shaping group dynamics. Studies have shown the positive effects of similarity of expectations [60, 35, 36], and how these shared expectations gradually develop as the team performs more tasks together and team members get to know each other better [74, 58].

The Team Mental Model (TMM) construct was introduced as a way of capturing the implicit coordination observed in effective teams and understanding how they operate in complex, dynamic context [60]. A team with well-developed TMM should have a common view of what is happening, what is likely to happen next, and why it is happening. Having a shared understanding of the situation permits the team to coordinate actions and adapt its behavior as a response to changes in task demands, leading to enhanced decision making and higher performance. In contrast, teamwork breakdowns or ambiguity with respect to who is responsible for specific tasks may lead to catastrophic results [34, 88].

TMM presents two main properties as measures of team functioning—similarity and accuracy. Similarity measures the degree to which understandings are alike

among individuals—for example, whether the team members see the team priorities in the same way. Accuracy measures to what extent the team's beliefs reflect the "true state of the world" [32]. Both properties are important for understanding team performance—for example, similarity combined with inaccuracy might cause a team to work efficiently, but towards the wrong goal.

These properties cannot be directly measured using our badges. However, disagreements may appear as a high level of turn-taking in meetings, with participants cutting each other off or frequently speaking at the same time.



# Chapter 3

## Hardware Design

### 3.1 Initial Requirements

The primary requirement of the system is to capture the conversational dynamics of a group by recording audio signals from the team members. The secondary requirement is to keep the system as simple as possible so it can be easily deployed at large events.

Prior studies conducted by the Human Dynamics group marked conversation dynamics as a key component affecting teams' performance [69, 89]. We therefore decided to focus this work on speaker detection, as this signal provides the most value and cannot be easily implemented by other means.

The main factor affecting the design of the board was simplicity, as it helps us to achieve several goals. By using only a small number of electronic components, all of which are widely available, we could keep the cost of the system low, rapidly prototype, and allow other researchers to reproduce the work more easily.

Additionally, the badge must have a small form factor (smaller than an average business card) and a low-profile, streamlined design so it can fit inside a standard plastic name tag. This obviates the need for designing and manufacturing a specialized enclosure, further lowering the cost of the badges.

These characteristics will also allow us to quickly deploy many of these badges at a large event, such as a public Hackathon or conference. At this type of event, the organizers have little time to interact with individual participants and ask them to

join a study, let alone perform an action such as installing a mobile application or filling out forms. We envision participants receiving a name tag already containing an active badge so that data could be collected from the moment they enter the event.

A low per-unit price and a simple enclosure make the study of these events more feasible, as the combined cost of badges and labor required for deployment is low. Another element in the cost calculation is the price of lost units—people are expected to be only weakly tied to the organizers of such events, and researchers should expect that some badges will be lost at each event.

## 3.2 Functional Description

The system consists of base stations and badges that communicate using Bluetooth low energy (BLE). An optional Wi-Fi access point enables easy access to the base stations. Figure 2-1 shows the components of the system.

The base station’s role is to keep the clocks of all badges synchronized, monitor their health and periodically pull data from each badge and store it locally on the base station. The badges’ role is to continuously collect speech data from each participant and store it until it get sent to the base station. When a badge is first activated, it will wait for the base station to connect and synchronize its internal clock.

A LP2985AIM5-2.0 low noise linear regulator drives the INMP401 breakout board to ensure a continuous, noise-free power supply for the microphone. The output of the INMP401 is connected to a passive low-pass filter with a cut-off frequency  $f_c = 312$ . The filtered output is connected to one of the Analog-to-Digital Converter (ADC) pins of the RFD22301 Bluetooth-enabled microcontroller (RFDuino). The microcontroller then samples the INMP401 microphone every 200ms for a period of 100ms and averages the values, creating five samples per second. Samples are stored on the SRAM of the RFD22301.

Due to the limited SRAM of the RFD22301, only 10 minutes of samples can be stored on the device. The base station connects to each badge before the samples buffer overflows and pulls the data over a BLE connection to ensure that no data is



Figure 3-1: A simple deployment that includes two badges, a base-station, and a portable Wi-Fi access point

lost.

### 3.3 Badge Design

The main goal in the design of the badge was to keep it simple, so it is easy and cheap to make in quantity and potentially instrument a large number of participants with badges. The final design has only a small number of components—a microcontroller with integrated Bluetooth Low-Energy transceiver, an off-the-shelf microphone breakout board with an amplifier, a coin cell battery holder, optional break-out programming header, a voltage regulator, a power switch, and several capacitors and status LEDs. These components are mounted on a 2 7/8" by 1 3/4" double-sided PCB board small enough to fit in a standard convention badge holder (Figure 3-2), or the badge can be simply worn as-is using a neck cord (Figure 3-3). Cost per unit

is \$38 for a small batch, not including assembly costs.

We wanted to allow other researchers who are interested in using this system to adopt it easily, and therefore used widely available components that are also easy to hand solder (see Appendix B). For the same reason, we decided to use Arduino [2], one of the most popular development environments for hardware projects.

Following techniques pioneered by the Sociometric Badges, badge captures conversation dynamics by comparing the volume observed by each of the badges in order to determine which participant spoke and when. The values are averaged in a way that obscures the actual content of the conversation. Chapter 4 describes the data collected by the badges and the way we analyze them in more detail.

Other methods for speaker diarization exist, but are usually domain specific and require recording the actual conversation [81, 59]. Besides the technical implications of recording the conversations (e.g. high bandwidth requirements for the transmission or stronger computational resources on the badge itself), people are typically reluctant to be recorded [25].

In its current form, the badge operates as a stand-alone data-acquiring unit that can save up to 10 minutes of data before transmitting it to a base-station. It can operate for up to 36 hours using a single CR2032 coin cell battery, due to the low-power design.

The next sections describe the different components used in the badge and the considerations for each one.

### **3.3.1 Bluetooth Low-Energy**

The core component of the badge is the RFD22301 (also known as RFduino), a 16MHz ARM Cortex-M0 microcontroller with integrated Bluetooth low energy (BLE), based on Nordic nRF51822 [10]. Compared to the other popular BLE-enabled microcontrollers that were available at the time, the RFD22301 offered the best combination of price, size and ease of use, together with Arduino support (see Table 3.1). Using an IC that packs both a microcontroller and a BLE transceiver, the RFD22301 further simplifies the hardware and software design of the badge.

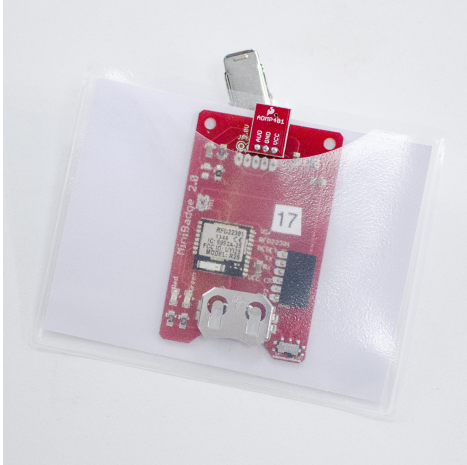


Figure 3-2: A badge carried inside a convention badge holder



Figure 3-3: A badge with a neck cord

Table 3.1: Popular BLE modules

| Module                      | Price   | Mounting | Size           | Integrated Antenna | Arduino support |
|-----------------------------|---------|----------|----------------|--------------------|-----------------|
| RFD22301 (RF-Duino)         | \$14.99 | SMD      | 0.6" x 0.6"    | Yes                | Yes             |
| RedBear BLE Nano            | \$17.90 | PTH      | 0.83" x 0.73"  | Yes                | Yes             |
| Punch Through Design LBM313 | \$10    | LGA      | 0.63" x 0.47"  | Yes                | No              |
| Punch Through Design Bean   | \$30    | PTH      | 1.75" x 0.875" | Yes                | Yes             |

BLE technology offers several advantages over ZigBee, ANT, Wi-Fi and traditional Bluetooth. As the name suggests, the BLE standard was developed to deliver ultra low-power wireless communication at a low cost [4]. Compared to other technologies, it provides excellent power-per-bit and a short set-up time [29, 77]. The RFD22301 specifications states a 12mA transmit/receive current, 4mA ARM CPU running current and ULP current of 4uA [10].

Another major advantage of BLE over other standards is that BLE has an increasingly ubiquitous presence on smartphones and laptops. This enables us to communicate with the badges using most existing platforms. Chapter 3.4 describes one such implementation using Raspberry Pi—a pocket-sized, single-board computer.

### 3.3.2 Microphone

A commercially available breakout board with a MEMS microphone and an amplifier is placed parallel to the main board, such that the microphone is facing the chest of the wearer. This design focuses the microphone and makes it more sensitive to the sounds generated by the wearer, protects the microphone, and creates a streamlined, low profile.

Compared to electret microphones (ECMs), microelectromechanical microphones (MEMS) provide better noise performance, device-to-device repeatability, and low power consumption, all in a very small package [55]. However, using MEMS devices involves a reflow soldering process that requires equipment not available to the average graduate student.

The Sparkfun INMP401 [12] breakout board was selected because it is one of the only off-the-shelf breakout boards available with a MEMS microphone, the Analog Devices ADMP401 [1]. This microphone has a flat frequency response from 100 Hz to 15 kHz. Since reflow soldering requires special equipment, using a pre-assembled board reduces prototyping time and makes it easier to reproduce this thesis.

The INMP401 also contains an amplifier with a gain of 67, and a DC offset circuit that offsets the output of the microphone around  $VCC/2$  voltage, producing a microcontroller-friendly signal.

### 3.3.3 Microphone Low-pass Filter

As explained in Chapter 4, when determining the speaker for each timestamp, we compare the sound levels captured by all badges. A passive low-pass filter was therefore added to reduce the noise from unwanted sources, which can mask the participant's voice.

The fundamental frequency for adult human speech ranges between 85 Hz and 255 Hz, with women having a higher mean fundamental frequency than men within this range ([80, 16]). For convenience, we set the cut-off frequency at  $f_c = 312$ .

A first order low-pass filter circuit consists of a resistor in series with a load, and

a capacitor in parallel with the load. The cutoff frequency can be calculated using the following equation:  $f_c = \frac{1}{2\pi RC}$ . By setting the cutoff frequency  $f_c = 312$  and the capacitor value to  $C = 0.1\mu F$ , we can determine the resistor value  $R = 5.1k\Omega$ . Because of the built-in 312 Hz cutoff frequency of the microphone, we receive a 100 - 312 Hz band.

### 3.3.4 Power Supply

The badge is powered by a single CR2032 3.0V coin cell battery for up to 36 hours. We chose to use this type of battery because of its small size and low price—as low as 30 cents per unit. This type of battery is widely available and easy to replace.

The alternative, rechargeable batteries, may be cost-effective in the long run. However, a large deployment with hundreds of badges would have a high initial cost—not only the price of the batteries themselves, but also the cost of battery chargers. In addition, each deployment would require a considerable amount of human resources to keep the batteries charged.

### 3.3.5 Microphone Voltage Regulator

The main board includes a low power LDO power regulator to ensure a continuous, noise-free, power supply for the microphone. The Texas Instruments LP2985AIM5-2.0 operates from +2.2V to +16V, and delivers a guaranteed 150mA low-noise current with a low-dropout at +2.0V[7].

The +2.0V supply voltage falls well within the required operational range of the INMP401 unit (1.5-3V), while the 2.2V supply voltage required to drive the LP2985AIM5-2.0 unit itself can still be delivered by a 3.0V coin cell battery.

## 3.4 Base Station

A base station is required for collecting the data from the badges. Since the badges have limited memory, we developed a base station that reads the data from each device

every couple of minutes. Additionally, the base station synchronizes its internal clock with an NTP server and sets the clock on each badge in order to ensure synchronized timestamps.

The base station is implemented using Raspberry Pi Model B+, a low-cost, low-power, pocket-sized, single-board computer with 512 MB RAM, 700 MHz single-core ARM processor and 4 USB ports [9]. We also included a Wi-Fi module for network connectivity and a BLE module to communicate with the badges. The base station code uses a combination of the official Linux Bluetooth stack, BlueZ [5], and BluePy, an open-source BLE library for Python [3].

The Raspberry Pi computers enabled us to create a low-cost, stand-alone system that is easy to deploy and simple to extend. For example, multiple base stations can be deployed in order to get good coverage if required. It is also possible to deploy several base stations and scan for devices in order to calculate an approximate location based on the RSSI values [23, 43].

A dedicated server code manages the communication with the badges. At startup, the server program connects to each badge and sets its internal clock so that the samples' timestamps are synchronized. It then queries the badges in a round-robin fashion and pulls a compact data structure with the latest data samples.

Each payload of data consists of the time stamp of the first sample, the sampling delay between samples, and a series of samples. Using these data, we can later assign the exact time stamp for each sample. The payload also contains the current voltage of the badge's battery for monitoring purposes.



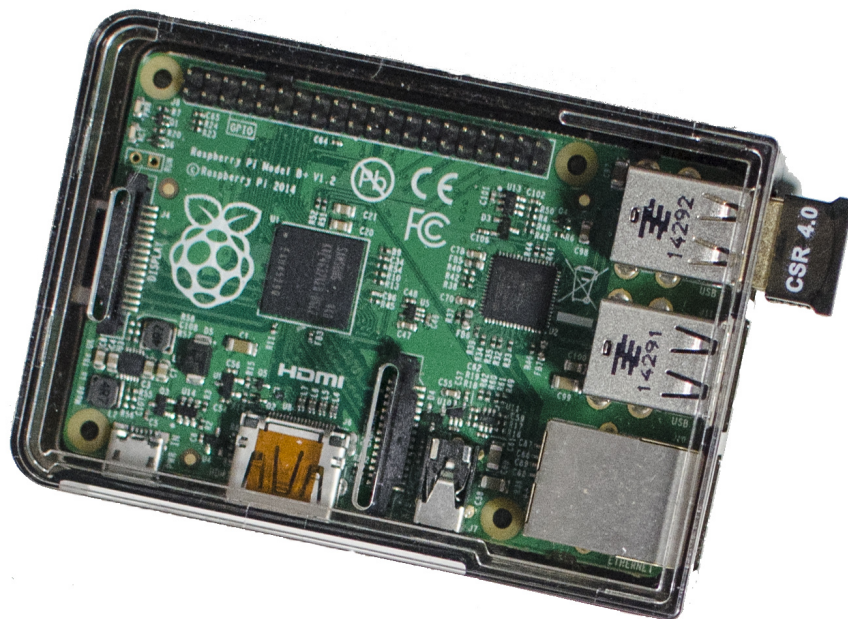


Figure 3-4: A Raspberry Pi base station



# Chapter 4

## Data Analysis

### 4.1 Determining Badge Volume

The badge creates a data point every 200 ms by averaging samples over a period of 100 ms. This allowed the microcontroller to be placed in an energy-saving mode for almost half the time. Sampling rate is kept at  $f_s > 800$  Hz, due to the 100 - 312 Hz frequency band created by the filters on the badge.

The ADC on the microcontroller converts the analog value read from the microphone to  $s(t)$ , a value between 0 and 1023 that corresponds to an input voltage of 0v and 3.6v respectively. The microphone breakout board DC offsets the signal to 1v, half of the operating voltage provided by the 2v voltage regulator.

Taking into account the center voltage of 1v and a range of values between 0v and 2v, the amplitude is calculated and then averaged to create a data point:

$$a(k) = |s(k) - 1023 \frac{1}{3.6}| \quad (4.1)$$

$$v(t) = \frac{1}{f_s T t} \sum_{k=1}^{f_s T t} a(k) \quad (4.2)$$

Where  $f_s T t$  is the number of samples used for creating a data point of time  $t$ . We then use a threshold and a moving average transformation on the data series to smooth the signal and remove noise.

## 4.2 Detecting the Active Speaker

The active speaker at each time interval is determined by comparing the amplitude recorded by all of the badges that took part in the conversation at time  $t$  and choosing the badge with the highest amplitude:

---

**Algorithm 1** Speaker selection algorithm

---

```
1: procedure ACTIVE_SPEAKER( $t$ ) ▷ time
2:    $activeSpeaker \leftarrow none$ 
3:    $maxAmp \leftarrow 0$ 
4:   for each badge  $n \in N$  do
5:     if  $v_n(t) > maxAmp$  and  $v_n(t) > 0$  then
6:        $activeSpeaker \leftarrow n$ 
7:        $maxAmp \leftarrow v_n(t)$ 
8:     end if
9:   end for
10:  return  $activeSpeaker$ 
11: end procedure
```

---

An example for the behavior of this algorithm can be seen in a dataset collected in one of our preliminary experiments. In these experiments, two participants were reading monologues and dialogues taken from the television sitcom Seinfeld. Figure 4-1 shows badge #4 reading a monologue between minute 31:15 and 32:00. The dataset was then processed using the method described in the previous section to produce Figure 4-2. In the two figures we can see activity on both badges, but since the amplitude from the participant wearing badge #4 is higher, the algorithm chooses him as the active speaker (Figure 4-3).

## 4.3 Computing Turn Taking

Turn taking is computed by counting the number of times the active speaker has changed, as described in Algorithm 2. Examples for the output of this algorithm can be found in Chapter 5.

Note that we assume that only one speaker is active at any given time. This may result in a high level of apparent turn taking when several people are talking

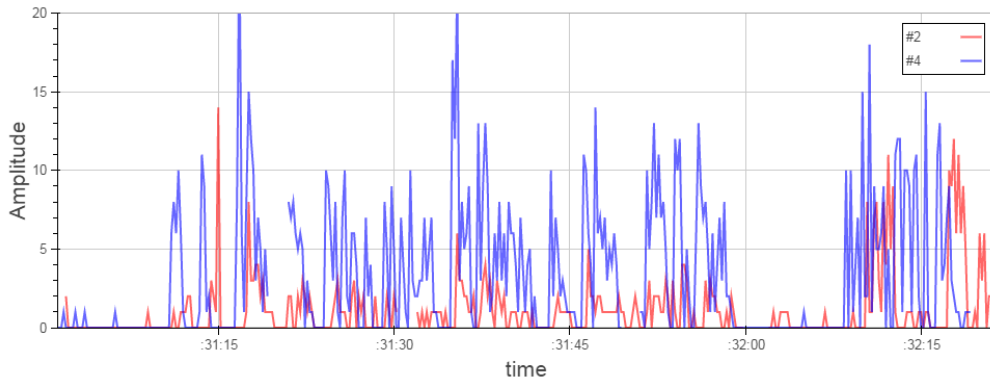


Figure 4-1: Raw vocalization samples from two badges. Badge #4 is reading a monologue between 31:15 and 32:08

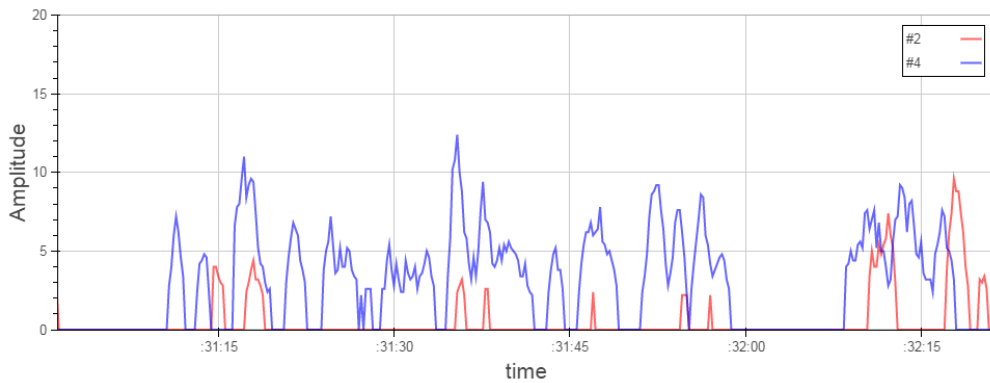


Figure 4-2: Monologue data after MA

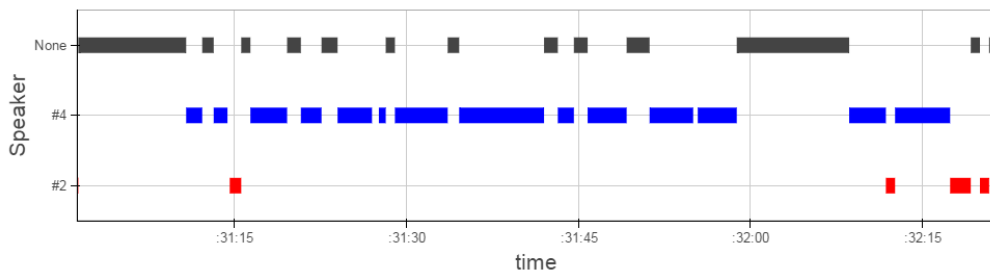


Figure 4-3: Speaker detection. Red shows badge #2 as the speaker. Blue shows badge #4 as the speaker. Black represents silence

---

**Algorithm 2** Turn Taking Count

---

```
1: procedure COUNT_TURN_TAKING( $t1, t2$ ) ▷ The time interval
2:    $count \leftarrow 0$ 
3:    $lastActive \leftarrow none$ 
4:   for  $t = t1$  to  $t2$  do
5:      $currentActive \leftarrow ACTIVE\_SPEAKER(t)$ 
6:     if  $currentActive \neq lastActive$  and  $currentActive \neq none$  then
7:        $count \leftarrow count + 1$ 
8:        $lastActive \leftarrow currentActive$ 
9:     end if
10:  end for
11:  return  $count$ 
12: end procedure
```

---

simultaneously.

# Chapter 5

## Badge Evaluation

We evaluated the badges using a series of experiments, with each experiment testing a different aspect, such as people’s reaction to the badges, the stability of the system, and whether the signal is rich enough to study team behavior during short events.

We first tested the badges in a short team-building exercise called "The Marshmallow Challenge"—an activity in which participants are asked to work together and build the tallest freestanding tower made of spaghetti sticks and tape that can hold a marshmallow at the top.

The results of these studies help us validate the functionality of the badges by showing that the observed behavior matches the data collected by our framework. Our analysis also suggests that the exercise itself can be used to study different factors affecting team performance, such as team composition and size.

We then conducted two Hackathon studies. The first, MIT Hacking Medicine, was primarily used for testing the stability of the data-collection platform over a two-day period, and testing the willingness of participants to wear it. Due to technical difficulties, keeping the system working properly took a considerable portion of our time, but we were still able to make general observations on the behavior of the two teams we monitored and compare it to the data collected by the badges.

For the second Hackathon, an internal corporate Hackathon, we focused on making detailed notes on the behavior of team members and the group as a whole so we could later look for these behaviors in the data and evaluate Hackathons as a research

setting.

The preliminary results from these Hackathons show that, as expected, the teams follow patterns and behaviors similar to the ones described in the teams literature.

## 5.1 The Marshmallow Challenge

The Spaghetti and Marshmallow challenge is an eighteen-minute team-building exercise in which participants are asked to very quickly develop a solution to an unexpected problem under strict constraints (e.g. time and materials).

Here, we show that by instrumenting the participants with badges, we can measure the contribution of each team member to the discussion, and observe how much time the team spends discussing the problem and how the energy of the discussion changes throughout the experiment.

Due to the short duration and the similarities between this event and the planning and execution phases in Hackathons, Marshmallow Challenge teams can be seen as the fruit-fly version of Hackathon teams. These characteristics enable us to test the effect of several factors on team efficacy, such as the effects of group size and the composition of skills on the dynamics and performance of teams, in a more controlled environment. The challenge also provides a simple, clear measure of success—the height of the tower constructed by the team.

### 5.1.1 Background

The Spaghetti and Marshmallow Challenge is a team-building exercise that was first introduced by Peter Skillman in 2006, and later popularized by Tom Wujec in his 2010 TED talk [92]. In this challenge, small teams are given 18 minutes to build the tallest freestanding structure that holds a single marshmallow on top.

Tom Wujec’s Marshmallow Challenge website [91] provides detailed instructions for the experiment: each team is provided with a kit that includes: scissors, 20 pieces of spaghetti, three feet of string, three feet of tape, and one marshmallow. When the countdown clock starts, they have exactly 18 minutes to build the tallest structure



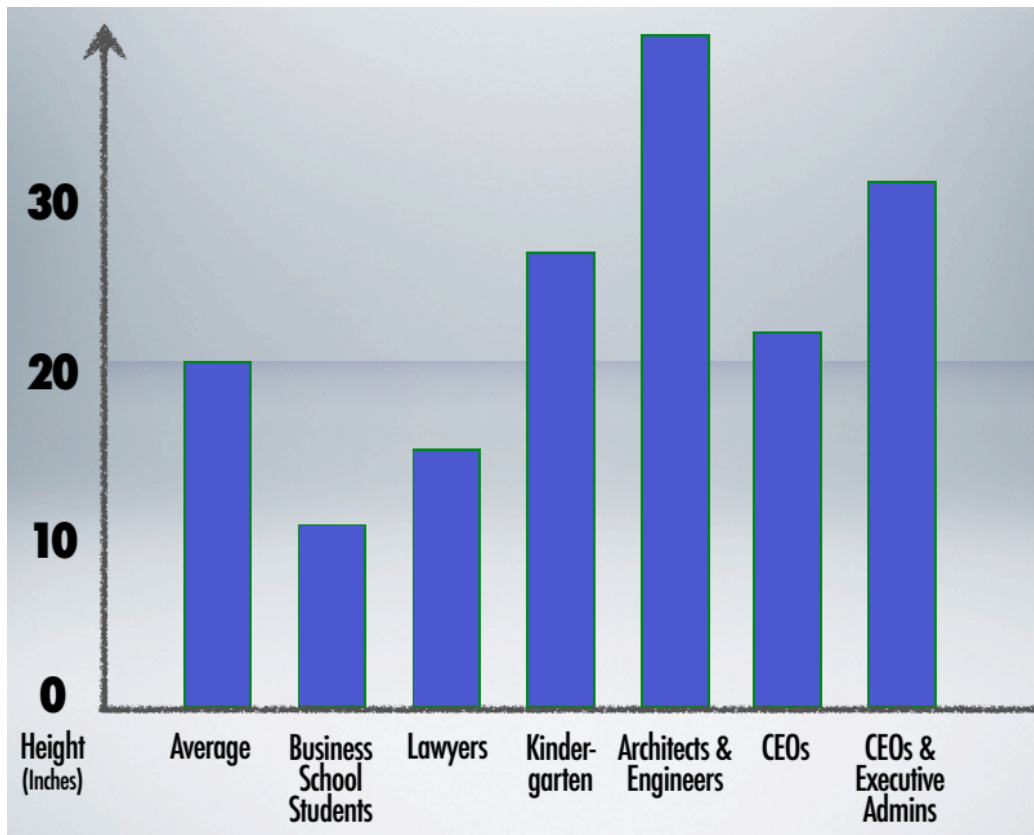


Figure 5-1: Performance of different types of teams. Source: Tom Wujec, TED April 2010

that can hold the marshmallow using the materials in the kit. A proctor reminds the groups of the remaining time and measures the height of all qualifying structures.

Wujec has conducted over 70 design workshops with engineers, executives designers, business school students, and even kindergartners and shared his findings about the performance of different types of teams and their style of collaboration. Table 5-1 summarizes the results from these experiments.

Interestingly, recent graduates of business schools tend to significantly underperform compared to the average, a finding that Wujec attributes to their training—discussing and conceiving the single perfect plan and executing it. They spent considerable time trying to establish dominance and plan, leaving only a few minutes to execute their idea. By the time they placed the marshmallow at the top of the tower and realized that it is not as light as it seems, it was too late.

In contrast, successful teams in this challenge showed the characteristics of a lean

startup—they learned by doing, rapidly prototyping and experimenting. They broke into smaller teams, worked in parallel and performed multiple iterations. Not surprisingly, kindergartners did well above the average. Instead of wasting precious minutes discussing who was going to be CEO, they started building and trying different shapes, gaining better understanding of the structural properties of spaghetti and the weight of the marshmallow.

Wujec also highlights the importance of diverse skills. Engineers and architects, through years of experience, obtained the specialized skills required for building tall structures using the materials provided in the experiment. Not surprisingly, the tallest stable towers were indeed built by engineers and architects. Facilitation skills matter as well—while teams of CEOs did better than average, putting an executive administrator in the team significantly improves its performance. Wujec hypothesizes that their facilitation skills help them manage the process and keep track of time. This result is consistent with other studies that showed that diversity of skills and information lead to better team performance.

### 5.1.2 Study Description

We recruited two teams of students and staff from MIT Media Lab and MIT Sloan School of Management for this experiment. Each team was given the kit and the instructions described on the official Marshmallow Challenge website [91].

In addition, each participant was asked to answer a series of short surveys that included demographic and professional background questions as well as the Ten-Item Personality Inventory (TIPI) [37] of the Big-Five personality dimensions. We then instrumented the participants with badges. A unique identifier was issued for each participant to link her survey data and badge data and protect her identity.

During the experiment, the remaining time was displayed on a tablet and announced by the proctor. The proctor also took notes describing the progress of the group, when it started trying the materials and when the marshmallow was first placed at the top of the tower. In contrast to the original setting of the Marshmallow Challenge, the teams could not observe the progress of other teams since only one

team participated in the experiment at a time.

### 5.1.3 Results

#### Team #1

Team #1 consisted of four participants from the same Media Lab research group and one participant from Sloan. The participants were allowed to introduce themselves before the experiment started.

At the very beginning, participant #15 suggested that since the marshmallow is heavy, they should start building with the marshmallow on top so that they could see how its weight is supported by the spaghetti. The team also agreed that they should split into two groups and maximize their time.

Data collected during the experiment show several stages in the team's work. Figure 5-2 and Figure 5-3 reveal that in the first five minutes of the experiment, the team mostly discussed the problem and possible solutions. In the next five minutes the amount of conversation was dramatically lower as the team started trying out different approaches.

Not shown in the data is the fact that shortly after the counter started, the team split into two smaller groups. At first the groups tried different approaches to the problem. Then, after a short group discussion for sharing insights and coordinating their efforts, two team members focused on the base of the structure, while the remaining three members were building the top part.

In the second half of the experiment we can see the amount of conversation increasing again as the team is running out of time, except a short period of silence just as the team put together the two parts of the structure. This team produced a 21" tower.

#### Team #2

Team #2 had a total of three participants from different programs in the MIT Sloan School of Management. They started discussing the problem and different strategies,

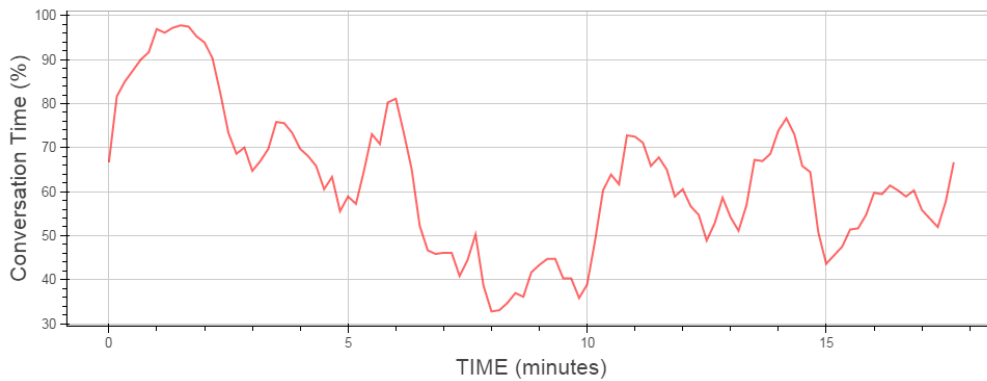


Figure 5-2: Team #1, percentage of conversation time per minute

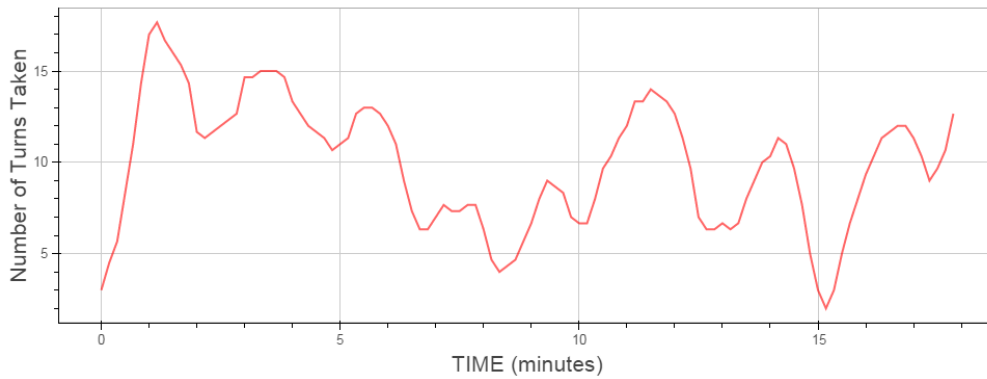


Figure 5-3: Team #1, number of turns taken per minute

and very quickly tried out the available materials—after three minutes they already had an untaped pyramid shape.

Once they agreed on a strategy (three pyramids with doubled spaghetti holding a single-spaghetti pyramid) they started taping the pieces of spaghetti together to create the tower. As seen in Figures 5-4 and 5-5, starting at the seventh minute, there were fewer conversations, and hence more building. They did not stop talking, though—during this time the participants kept discussing what needed to be done and how.

After the proctor gave the final five-minute warning, the group members realized how little time they had left, and had a short planning session followed by a burst of work. Looking at the last two minutes of the data from this team (Figure 5-4), however, one can see that as the team was running out of time, the team members

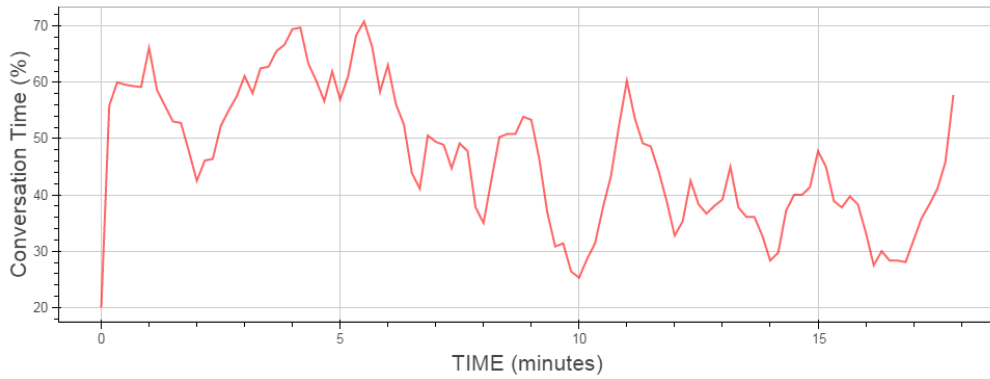


Figure 5-4: Team #2, percentage of conversation time per minute

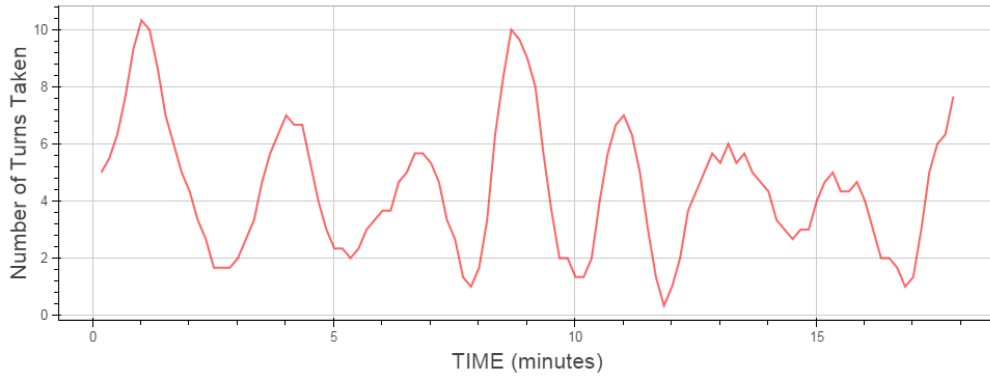


Figure 5-5: Team #2, number of turns taken per minute

spoke more to each other in order to coordinate their work. Team #2 built a very stable, 19.5" tower.

Interestingly, Figure 5-7 shows the participant wearing badge #16 in team #2 was clearly a more active speaker than the other team members. She was, in fact, dominating dominating the conversation. One would expect participant #15 in team #2 to dominate the conversation in his team, though—he is the former CEO of a small company, an MBA program graduate, and, based on the Ten-Item Personality Inventory, shows a higher level of extroversion than the other team members. However, he did not try to manage the team here. One possible explanation is that he has enough experience to let people do their work and only direct them when necessary.

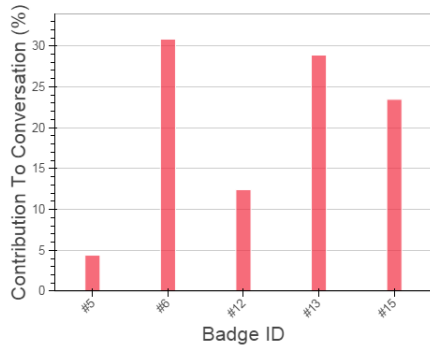


Figure 5-6: Team #1, Contribution of each badge to the total conversation time (%)

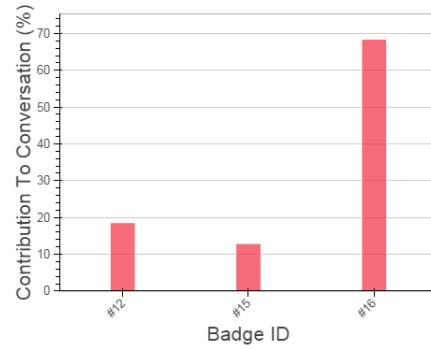


Figure 5-7: Team #2, Contribution of each badge to the total conversation time (%)

## Comparison

The comparison of the two teams reveals several differences in their behavior. While both teams showed an increase in turn taking towards the end of the experiment, it did so sooner in team #1. This may indicate a greater need for collaboration in team #1, especially since the team broke into smaller groups—participants #6 and #13 worked separately from the rest of the team and constructed the base of the structure while the rest of the team worked on the top part. Further experiments are required in order to determine whether this behavior is an example of a large team that was able to efficiently leverage its size by breaking into small independent working units after the initial discussion [27, 40, 48].

Prior studies have found the level of equality in contribution to the discussion to have a strong relationship with performance, as it represents equal opportunities to share information and opinions [70, 62]. Figures 5-6 and 5-7 show how the teams differ in the contribution their members made to the conversation. As observed by the proctor, the participant wearing badge #16 in team #2 was clearly a more active speaker compared to the other team members, and was in fact dominating the conversation. The conversation in team #1 seems more balanced. Particularly, the two participants who worked on the base (badges #6 and #13) had an almost identical contribution to the discussion.

## 5.2 Hackathons

### 5.2.1 Background

Hackathons provide an excellent bridge between the laboratory and the corporate setting, where project-based work is so important for innovative outcomes. In a Hackathon, groups of people come together for a very short burst of creative work, and are expected to go through the full range of tasks and phases composing the innovation process.

In this setting, people have a high incentive to be creative; after all, that is why they are there. This creativity can be quantified, based on scores provided by the judges, or, in a corporate Hackathon, how many resources are eventually allocated to the idea. Corporate Hackathons in particular enable us to study team behavior a setting where team members have a better understanding of why they are engaged in the task, and have an affiliation with it.

This section mainly serves as a proof of concept. It tests the feasibility of using the badges for capturing data from long, dynamic events. We then demonstrate how the rich signal captured by these devices can be used to study the changes in the behavior of teams throughout an event and identify different stages in the teams' life cycles.

We conducted two studies in different Hackathons. The first Hackathon was a large scale event that invited the public to help improve medical care. Here, we wanted to test participants' willingness to wear the badges for two days, and collect basic vocalization data see how different stages of the event are represented in the data.

The event also presented us with an opportunity to test the stability of the system in a dynamic, two-day event. Due to technical difficulties and the time it took us to recruit two teams for the study, only limited qualified data is available from this study.

The second Hackathon was a corporate event that brought together employees from different departments in order to improve customer-facing services. At this

point, we already had a stable system that required minimal maintenance, and we could therefore take notes and collect more detailed qualitative data.

### 5.2.2 MIT Hacking Medicine—Grand Hack 2015

Organized by MIT Hacking Medicine [22, 30] and held at the MIT Media Lab, the Grand Hack 2015 had a turnout of more than 450 participants from 19 states and eight countries, making it one of the largest healthcare Hackathons in the world [6]. A total of 80 teams competed for over \$13,000 in cash prizes as well as incubation and mentorship opportunities. Now in its second year, the Grand Hack is MIT Hacking Medicine's flagship event to innovate in healthcare by bringing the entire spectrum of healthcare stakeholders together in a single Hackathon.

The Hackathon opened with a welcome session and a series of presentations explaining the goals of the Hackathon, after which the event broke off into the four separate tracks—Global Health, Primary Care, Telehealth, and Wearables. Each participant was assigned to one of these tracks beforehand, based on his or her registration information.

In each track, participants pitched their "pain point"—a problem in healthcare that they wanted to solve. In many cases, these pitches reflected personal problems that they or their families were facing.

As the participants walked around, introduced themselves, discussed the problems and started forming teams and hacking their ideas, we approached several groups and requested that they participate in the experiment. Having only a limited number of badges available, we monitored the behavior of two teams. However, one team produced largely unusable data due to technical difficulties.

#### **Team #1**

The first team included six participants with a mixture of backgrounds—clinicians, economists, engineers and designers. Two of the participants were females and the rest males, ages 25 - 45. The team was working on a novel mobile application that



aggregates data from a patient's glucose monitor and activity tracker to help change his or her behavior and improve his or her wellbeing.

After the team agreed to participate in the study, each participant filled the surveys and received a badge. The team continued to wear the badges until the final presentation took place on the last day of the Hackathon. Badges were returned to the researchers before participants left the event in the evening, and were given back to the participants the following morning.

In general, the team worked together around a single, round table most of the time, excluding meals, short workshops given by the sponsors of the track (Jawbone and Microsoft), and similar activities.

Based on our observations, the participant wearing badge #7 seemed to be the decision maker in this team—when we first approached the team, the other team members were looking at her, waiting for her answer. The data collected by the badges indeed indicates that during the first day of the event, her contribution to the conversation (Figure 5-12) was about twice as high as most of the other participants, with the exception of participant #6.

The second day of data reveals a more balanced contribution to the conversations. Participants #6 and #7 no longer dominate the conversation. In fact, the contribution of all participants is very similar. Note that participant #2 was absent during most of the second day, and therefore no data is available from her badge.

Looking at the conversation patterns of the first day in Figures 5-8 and 5-9, we observe several interesting behaviors. The team had repeating cycles in which the conversation builds up and then fades. Some of this behavior can be attributed to specific events that we observed during the event. For example, we can see the conversation building up at 4:30 pm when the team returned from an organized activity, and fading again around 5:00 pm when dinner was served. The conversation built up again as the team members started to return from dinner and faded later towards the end of the first day.

A longer, more interesting example occurred on the second day (Figures 5-10 and 5-11)—both the turn taking and discussion time started to rise at 10:40 am and

came to a halt by 12:20 pm. At that time, the team members left the table and went to practice their pitch with the mentors of this track. Therefore, the pattern of conversation during this surge of discussions most likely indicates how the team performs under intense time pressure. Not surprisingly, another peak of activity occurred when the team returned to the table and discussed the issues raised by the mentors. It is worth noting that the lack of discussions during the practice time is due to technical limitations—the team moved outside the range of the base station for a long period of time, leading to data loss.

This Hackathon also presented an unexpected opportunity to measure the effect of an intervention on the behavior of the group. At about 4:00 pm of day 1 (Figure 5-8), all participants were invited to join a dancing activity meant to energize them. What appears to be a spike in conversation time is in fact noise created by the loud music played during the dance session.

It is difficult to tell what was the effect of this intervention on the behavior of the team. The conversation in team #1 started to build up over a period of 30 minutes prior to the activity, and therefore the activity might have cut off an important discussion. We can then see another discussion building up right after the activity ended—only to be cut off by dinner. While a more comprehensive study is required, the organizers might consider moving the dancing activity or dinner to different times, or better yet, determine the best time based on real-time data obtained by the badges.

At 2:00 pm of the second day, all of the teams stopped working on their presentations, gathered in the main conference rooms and started giving their final presentations. At that time, the participants gave back their badges.

## **Team #2**

The second team had nine participants, eight of whom agreed to wear the badges. The team worked in a very methodical way—each time the team assembled, the team members took turns explaining the status and giving their opinions. However, the team split into smaller teams working in different locations and making it difficult to collect data continuously since the participants spent long periods of time outside the

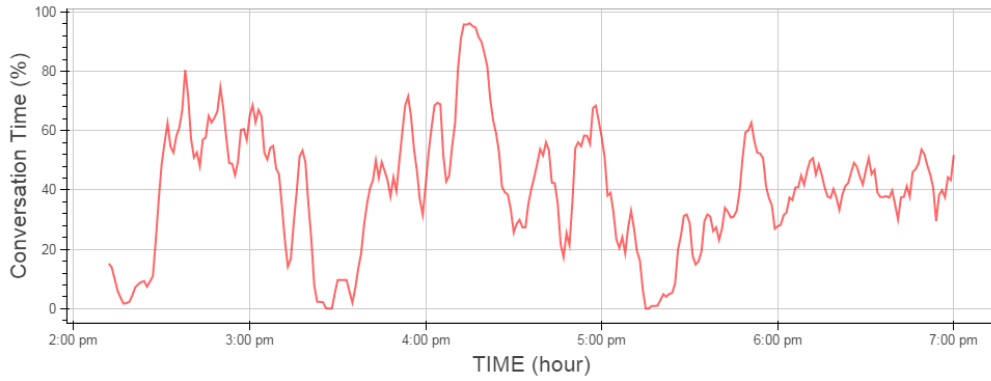


Figure 5-8: Team #1, Percentage of conversation time per minute, day 1

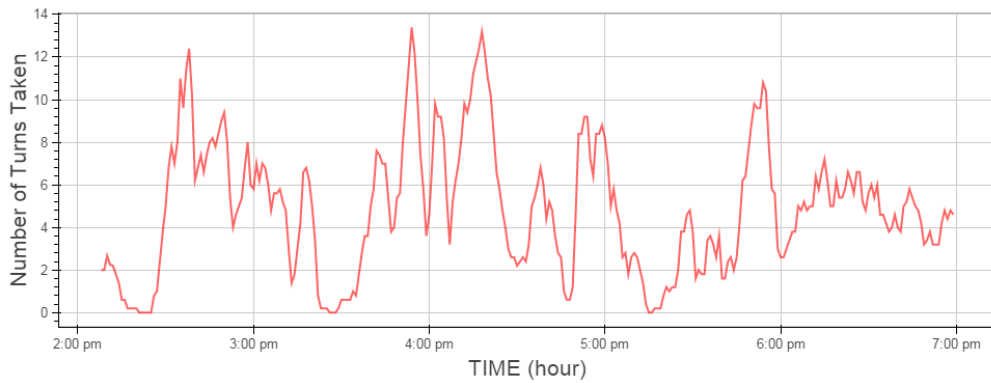


Figure 5-9: Team #1, Number of turns taken per minute, day 1

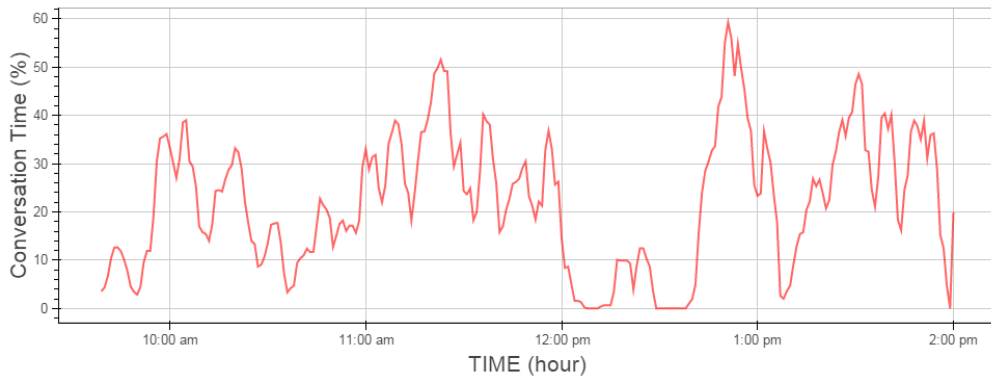


Figure 5-10: Team #1, Percentage of conversation time per minute, day 2

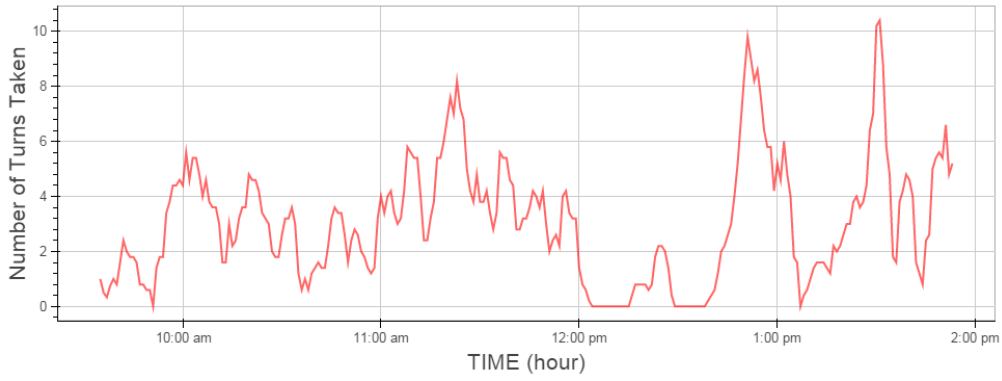


Figure 5-11: Team #1, Number of turns taken per minute, day 2

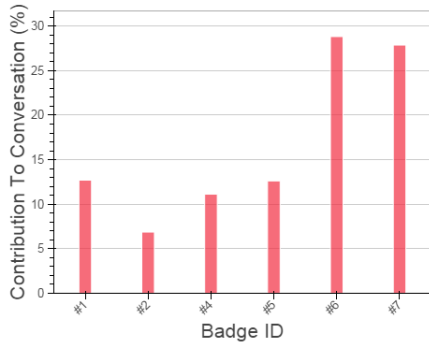


Figure 5-12: Team #1, Contribution of each badge to the total conversation time (%), day 1

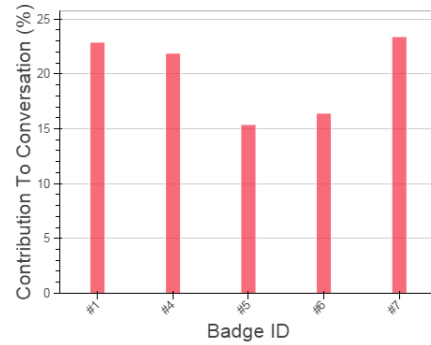


Figure 5-13: Team #1, Contribution of each badge to the total conversation time (%), day 2

range of the base station. Therefore, the dataset for this team is incomplete.

Figures 5-14 and 5-15 show the number of samples obtained from each badge in every minute of the two-day event. This information can be treated as a presence detector, telling us which participant was present around the table and when. If all participants are present, we can assume that they gathered for a meeting. Note that we only recruited this team at 3:00 pm, and therefore earlier data is not available.

The data reveal that team #2 did not hold many coordination meetings. In fact, it seems that they only met at the beginning and end of each day. While the team successfully split into smaller teams so it could be more efficient, this finding might indicate a lack of coordination between the sub-teams, since face-to-face communication has been shown to be the most valuable form of communication [69]. The ability to coordinate the efforts of all working units is essential for reducing uncertainties and generating trust, and eventually affects the performance of the team [86].

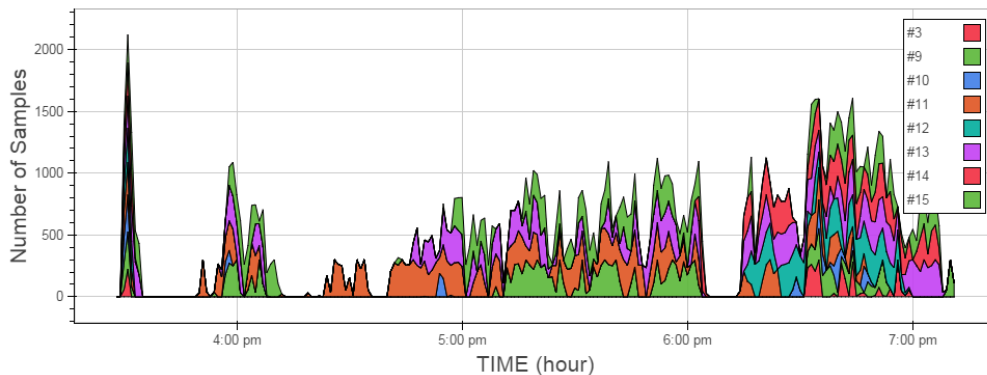


Figure 5-14: Team #2, number of samples collected by each badge during the first day

### 5.2.3 Corporate Hackathon By a Multinational Pharmaceutical Company

The next Hackathon we studied brought together employees and contractors from different departments of a pharmaceutical company, in order to improve the digital process of a customer-facing service. The company organizes an average of one event

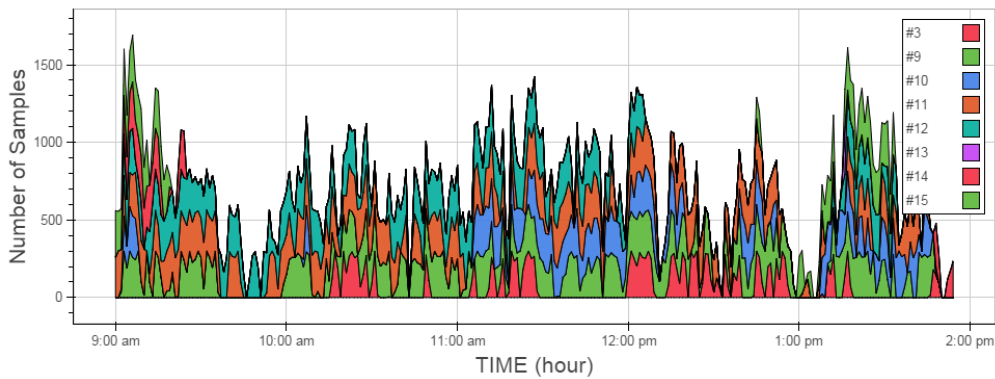


Figure 5-15: Team #2, number of samples collected by each badge during the second day

each month. Participants come to the table with a focused scope, based on pre-work meetings that take place a few weeks prior to the event.

As opposed to other Hackathons, these Hackathons are not competitive. Here, the entire team works as a single unit to define the problem, discuss how to solve it and then work together to prototype a solution.

The teams are typically cross-functional and include marketers, sales representatives, regulatory colleagues, a designer, several developers, a facilitator, and a subject matter expert.

On the first day, the organizers usually introduce the challenge and then the team members discuss the competitive landscape, go over insights from the pre-work meetings and review the ideas participants came up with during the discussion. After lunch the technical group begins working on the ideas, while the challenge owners and other business unit representatives are dismissed. The next day there is one checkpoint around lunchtime where the business unit representative comes back and sees the progress made on the prototype. A final presentation of the solution is demoed at the end of the day.

This specific Hackathon took place in the middle of the week in the company headquarters. The team consisted of about 12 participants, with several others occasionally joining the meetings. For some of them it was the first time participating in such an event. An offshore development team supported the implementation efforts,

especially at night.

The next sections provide a detailed description based on the notes taken by the author. They cover the various stages of the Hackathon and explain how the group dynamics are reflected in the data collected with the badges.

### **Stage 1—Challenge Presentation and Idea Generation**

This stage took place on the first day, between 9:00 am and 12:45 pm, and included an introduction to the concepts and goals of the Hackathon, followed by a discussion about the possible solutions. Figures 5-16, 5-17 and 5-18 show how the dynamics of the discussions changed throughout this stage, and will be explained in more detail shortly.

The event started with an introduction, in which the organizers (wearing badges #11 and #9) explained what a Hackathon is and what the goals were for the next two days. Afterward, at 9:20 am, the author of the thesis introduced the research project and handed out badges and surveys to the Hackathon participants.

Shortly after that, the participants paired up with other attendees whom they had not met before for a ice-breaking exercise—for half a minute, they were asked to engage in small talk while drawing portraits of each other.

Since the data collection system was not operational yet, no data was captured on this exercise. Even so, it still presents a potential intervention that can be studied in future research. Does it have a positive effect on the overall performance of the team? How long does the effect last? Do participants who complete the exercise interact significantly more intensely, compared to random pairs, friends and acquaintances?

From 9:30 am until 10:15 am, the corporate sponsor of the event (wearing badge #6) led a discussion about current solutions the company provides and possible improvements. The participants shared their knowledge, asked questions and got to know each others' specialties. Figure 5-18 shows that badge #6 was very active, especially at the beginning of this time period, as expected.

At 10:15 am the group took a 15-minute break. During that time, some participants left the room and some engaged in one-on-one conversations, as suggested by

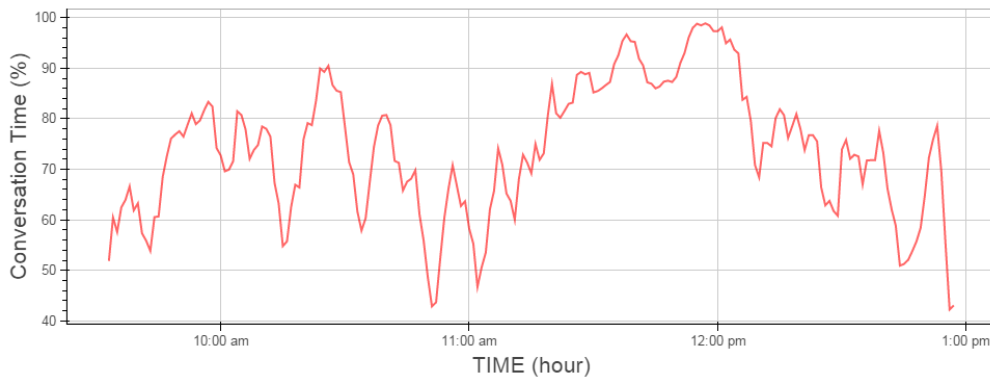


Figure 5-16: Percentage of conversation time per minute. Corporate Hackathon, Day 1, 9:30 am - 1:00 pm

Figure 5-16. However, looking at the turn taking (Figure 5-17), we can see one of the limitations of the badges—they cannot distinguish between different conversations that take place at the same time. Since we assume that only one person speaks at any given moment and identify the speaker by looking for the badge with the highest volume, having multiple simultaneous speakers registers as turn taking.

After the break, the group continued the discussion until 11:08 am. Figures 5-16 and 5-17 show that group was having an active, yet civilized and polite discussion—while the percent of conversational time was high, the amount of turn taking was low. Even though the group did not agree beforehand on a method of discussion (e.g. round-robin) they did not cut each other off very frequently. These results match the author’s observations and notes from this event.

A notable characteristic of these discussions is the opportunity they provided for participants to introduce themselves and contribute information that revealed their expertise. The familiarity and trust that the discussions facilitated may have enabled the participants to coordinate their next activities better [39, 73]. Ice-breaking activities, such the one the group practiced at the beginning of the event, are also likely to help the participants get to know each other better.

At 11:08 am, participant #11 summarized the ideas that the group discussed and explained the next activity. Figure 5-18 shows that this presentation had unique characteristics—she spoke almost uninterrupted for a long period of time.



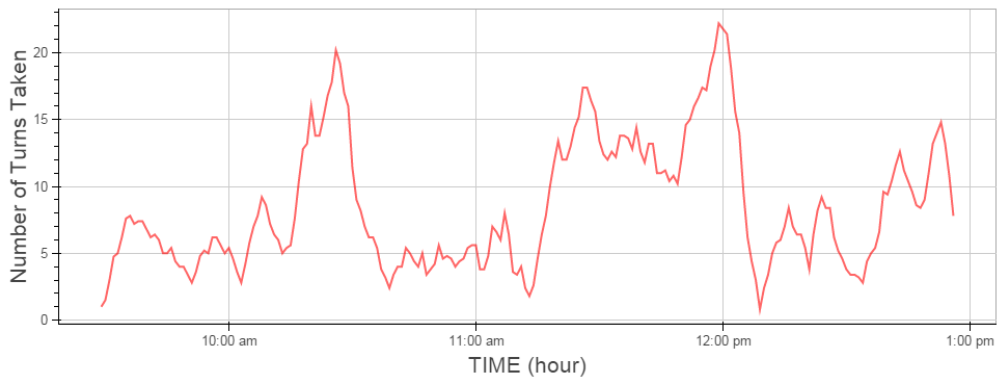


Figure 5-17: Number of turns taken per minute. Corporate Hackathon, Day 1, 9:30 am - 1:00 pm

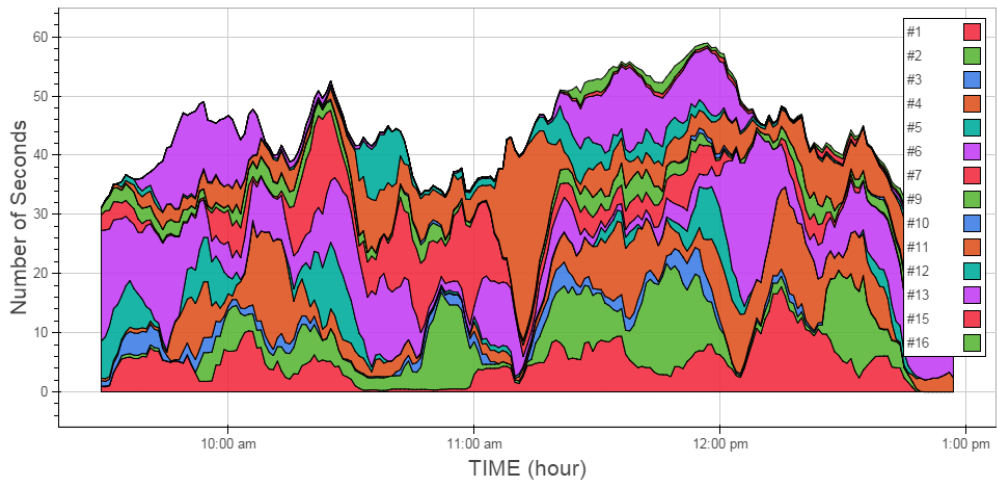


Figure 5-18: Contribution (in seconds) to the conversation, by participant. Corporate Hackathon, Day 1, 9:30 am - 1:00 pm. Values are averaged using a 10-minute sliding window in order to create a smooth, easy-to-read plot

In that activity, the participants were randomly assigned to two groups, located at the opposite corners of the conference room. We will refer to the groups as group #1 and group #2. Each group was randomly assigned a set of features from the list created during the previous stage, and it was up to the group to decide what they would focus on based on what they found most exciting and interesting. The participants gathered around a stand with large sheets of paper on which they could sketch their ideas. At the beginning of this activity, two new engineers (badges #15 and #16) joined the event and were assigned to group #1.

As suggested by the literature [27], by breaking into smaller groups and parallelizing the discussions, the participants were able to iterate on the ideas faster and go over more ideas in a short period of time. The participants were given instructions not to judge an idea, and instead of brainstorming and pushing new ideas, they were asked to improve an idea that someone had suggested and extend on that. The organizer explained this rationale as an alternative to "ending up with a board full of ideas".

The author mostly accompanied group #1 to take notes, but also tried to monitor group #2 in order to compare the behaviors of the two groups and see how the different composition and settings affected the dynamics. Figures 5-19, 5-20, 5-21 and 5-22 zoom into the time period between 11:15 am and 12:00 pm to show the difference between the two groups.

Based on our observations, the conversation in group #1 was more lively and active compared to the second group. These observations are also supported by the data collected from our badges—both the level of turn taking and the conversation time were higher in group #1 (Figures 5-19 and 5-20). In fact, this group was more engaged compared to any of the early discussions (except breaks). Considering that active discussion and information sharing are some of the goals in this stage, understanding the cause of these differences will help improve the productivity in future events.

One possible explanation is the difference in group size. Group #1 was larger than group #2—eight people compared six. With more people in the group, it is more likely to contain more opinions and have more relevant information to share [27].

However, group #1 was obviously smaller than the super-group containing group #1 and group #2, and yet it was able to engage in a faster, more active, discussion. Future studies should look into these mixed results.

The structure of this event may be an example for an efficient compromise—an orientation session with knowledge sharing that includes all the participants and helps create a common language and understanding [26, 36], followed by breakout sessions that utilize the power of smaller groups [27]. The increased turn taking after the participants returned to the table for a joint discussion at 12:05 pm, compared to the levels of turn taking before the breakout session, implies that these sessions have a noticeable effect on the group dynamics.

Another possible explanation for the differences between the groups is that the space available for group #1 was smaller, forcing the participants to stand closer to each other, and thus creating a sense of urgency. Having the participants of the breakout sessions stand, rather than sit around the table, may have contributed to this sense of urgency, the same way that stand-up meetings lead to a more focused, higher-speed discussion [18].

The two new developers who joined the group, wearing badges #15 and #16, were not very talkative during this activity (Figure 5-21). While this observation might indicate that a certain period of time is required for new members of a group to establish working relationships [73, 63, 39], Figure 5-22 reveals that participant #10 did not take an active part in his group discussion even though he joined the event at the very beginning. Alternative explanations might be specific personality traits that prevented these participants from joining the discussion, or a lack of information relevant to the problem.

At 12:05 am, the participants got back together, sat around the table and discussed the insights and ideas brought up by the groups. Figure 5-18 reveals the active participants in this part of the discussion—participants #1 and #2, who presented the results of their groups' discussions; participant #4, who manages the development team; and the organizers and main stakeholders of the event, participants #6 and #11. This part of the event ended at 12:45 pm.

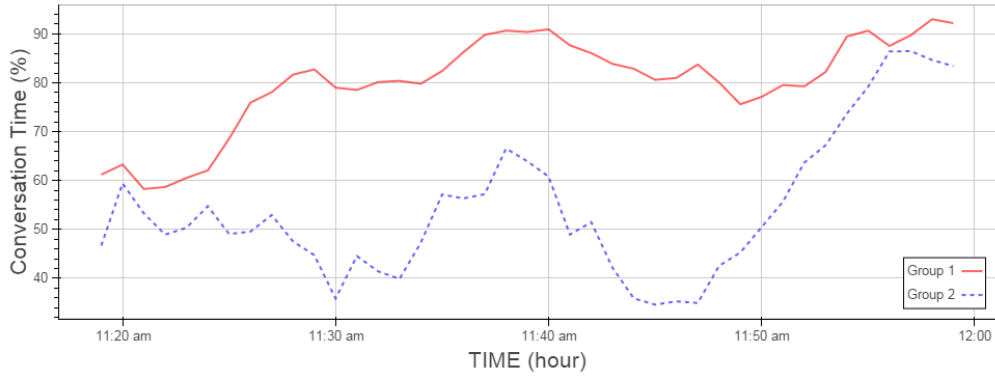


Figure 5-19: Percentage of conversation time per minute. Corporate Hackathon, Day 1, 11:15 am - 12:00 pm

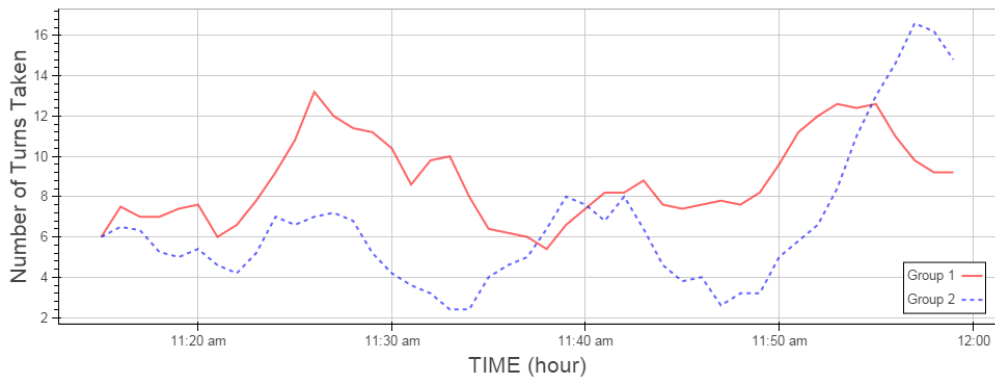


Figure 5-20: Number of turns taken per minute. Corporate Hackathon, Day 1, 11:15 am - 12:00 pm

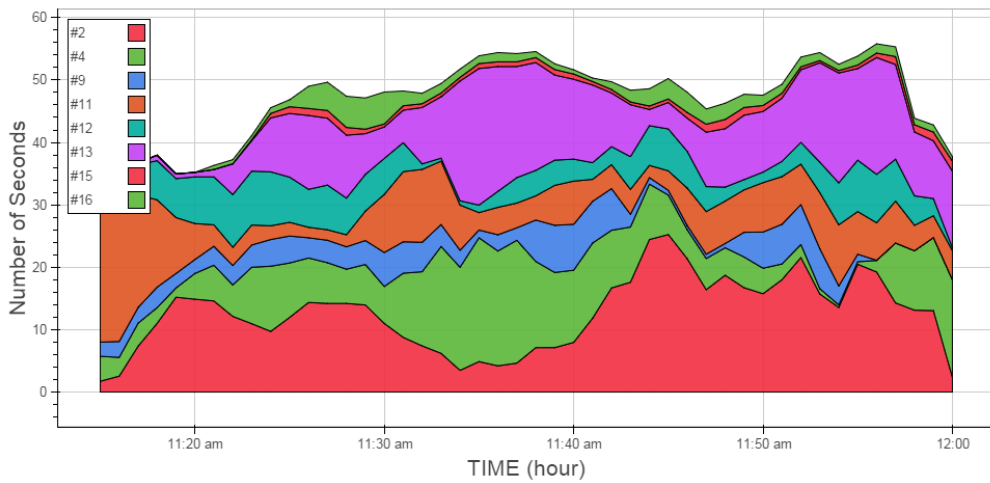


Figure 5-21: Group #1, Contribution (in seconds) to the conversation, by participant. Corporate Hackathon, Day 1, 11:15 am - 12:00 pm. Values are averaged using a five-minute sliding window in order to create a smooth, easy-to-read plot

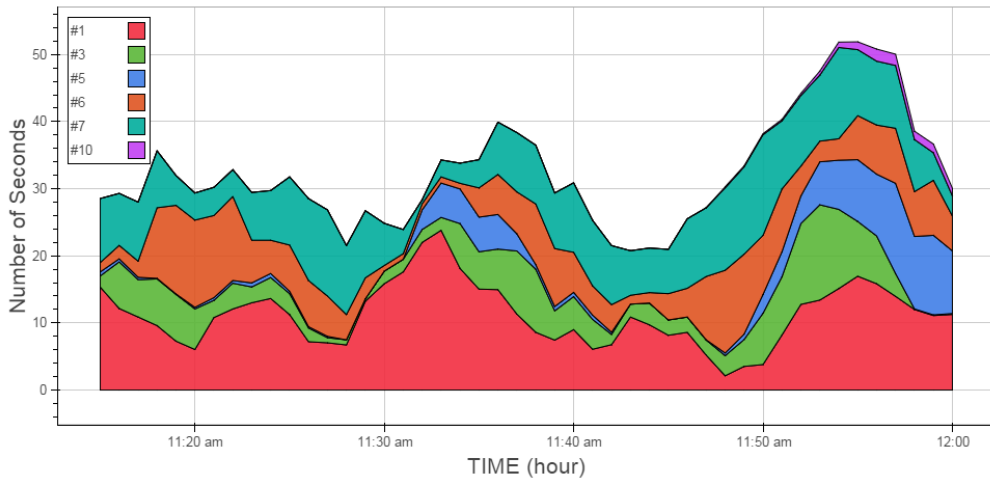


Figure 5-22: Group #2, Contribution (in seconds) to the conversation, by participant. Corporate Hackathon, Day 1, 11:15 am - 12:00 pm. Values are averaged using a replacedfive-minutefive minutes sliding window in order to create a smooth, easy-to-read plot

Figure 5-18 also reveals the contribution of the expert participants to different parts of the discussion. Participant #4, the head of the development team, was mostly engaged in technical discussions—for example, at the beginning of the event he was asked to give an overview of existing solutions. Then, as the conversation shifted from the user experience and how it can be improved to the details of implementation, he became more and more involved.

On the other end of the spectrum, participant #10 did not engage in the discussions at this stage. In fact, as we will show in the following sections, he was hardly involved at any stage of this event. We attribute this behavior to his role in the event—as a designer, he was less familiar with the specific requirements and concepts, and therefore did not have relevant information to share. Most of his interaction with the group consisted of short conversations during the implementation stage.

## **Stage 2—Implementation and Refinements**

This stage took part on the first day between 1:00 pm and 5:30 pm and on the second day between 9:00 am and 4:45 pm. During this stage, all the participants who were not involved in implementing the prototype left the room and only came back for status updates and the final discussion. The remaining participants sat around the table, first discussing implementation details, and later working on the solution itself in three smaller groups—back-end development, website design, and mobile application design.

At 1:30 pm the group started discussing how to prototype the solutions discussed earlier by the entire team. Not surprisingly, the manager of the development team (#4) and his team members (#15 and #16) were very active during the conversation as they provided more technical details and answered questions (Figure 5-25). The organizers (#9 and #11) were active throughout the discussion as well.

After discussing the timing and the next status meetings, the group agreed to keep working in the same room and prepare quick mock-ups for the 4:00 pm status meeting. Similarly to the teams in the other Hackathon and in the Marshmallow

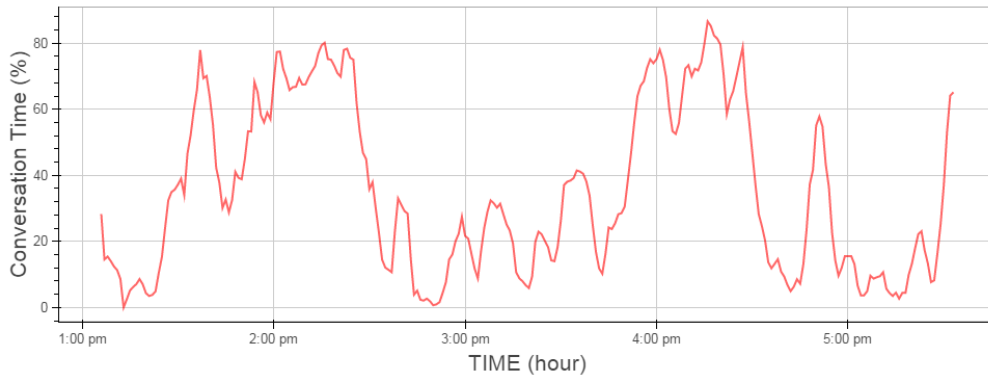


Figure 5-23: Percentage of conversation time per minute. Corporate Hackathon, Day 1, 1:00 pm - 5:30 pm

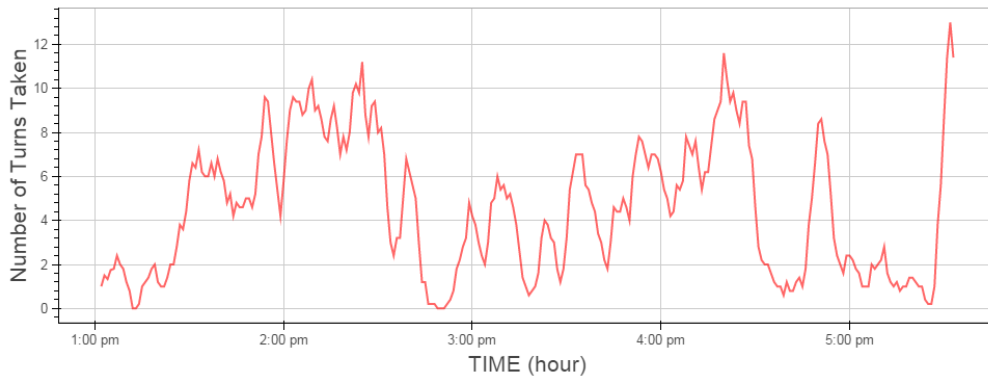


Figure 5-24: Number of turns taken per minute. Corporate Hackathon, Day 1, 1:00 pm - 5:30 pm

Challenge, we observed an increase in conversation time as the deadline got closer (Figure 5-23).

At 4:00 pm, participant #6, the corporate sponsor of the event, and participant #13 joined for a status report. The participants updated them on their progress and showed a mock-up of the new mobile application. During this presentation, the newly joined participants asked questions and provided input on the plan. The contribution of each participant is shown by Figure 5-25. As the corresponding areas in the chart suggest, participants #6 and #13 were very active during this discussion.

When the status update ended, there were several short discussions among the participants. Afterwards, most of them left to work from their own offices or get back

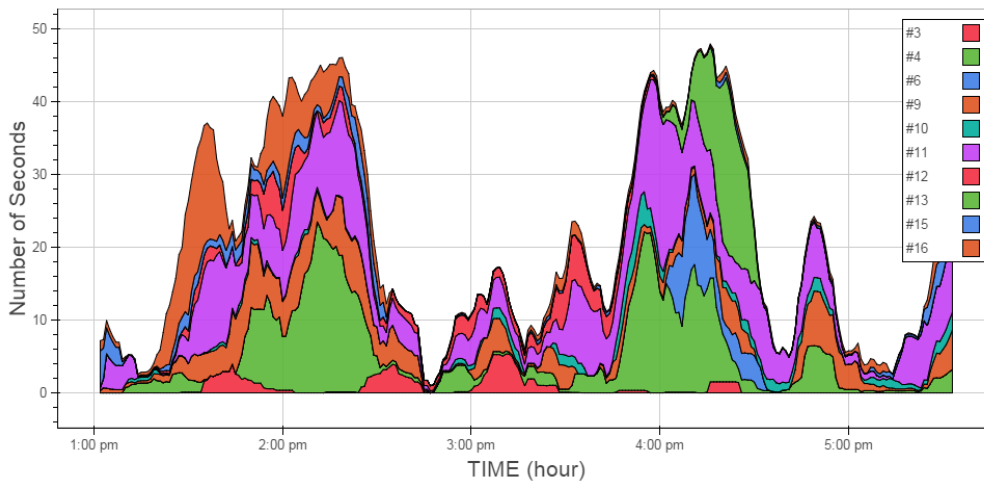


Figure 5-25: Contribution (in seconds) to the conversation, by participant. Corporate Hackathon, Day 1, 1:00 pm - 5:30 pm

to their usual work.

By 4:35 pm only four participants were left in the room. They working on the project, occasionally having short discussions. Then, at 5:30 pm, everyone packed their belongings and met some of the other participants for a dinner in a nearby restaurant.

The next morning, a small part of the group reconvened at 9:00 am and kept working on the implementation. The development manager gave everyone a quick status update, and they learned that an offshore team worked on the project throughout the night.

The group members were eager to see the prototype and prepare for the new status update that was scheduled for 12:00 pm, but because the developers did not arrive until 11:40 am they had to push the deadline to 12:30 pm.

Figures 5-26 and 5-27 clearly show the discussion building up again when the developers joined the rest of the group. Figures 5-28 provides additional information on the conversation dynamics and the active speakers in this discussion. Note that the manager of the development team (participant #4) hardly participated in the conversation. It is also worth mentioning that even though participant #16 was not wearing his badge during the discussion and instead left it on the table in front of



him, the badge proved to be sensitive enough to register when he was talking.

At 12:30 pm more participants came in, caught up with their colleagues and made some final preparations for the meeting. One of the participants joined remotely using the conference room phone. As before, small talk registers as a high level of turn-taking (Figure 5-27) since we cannot differentiate between conversations that take place at the same time.

Participant #6 then asked if everyone was ready, and began leading the status update at 12:45 pm. Most of the participants were either listening or taking active part in the conversation, except participant #16 who seemed to be busy finishing up the implementation.

After a short discussion about the compatibility of the rest of the websites owned by the corporation, participants #6 and #9 showed a mock-up of the website and the mobile application. At this point, participant #9 connected his mobile phone to the projector and gave a detailed presentation of the web application.

At 1:12 pm, participant #6 congratulated everyone on a job well done, and expressed her excitement about seeing the results and being able to show them to other partners in the company to facilitate a more fruitful discussion. The group then talked about additional changes to the user interface and agreed to meet again at 4:00 pm for a last update.

The sharp drop in conversation time and contribution marks the time when most of the participants left—1:30 pm. Those who remained discussed the feedback they had received and decided how to best use the time they had left.

When the group met again for the last time at 4:00 pm, we again see a lot of small talk before the main discussion starts. Then, if we look closely at Figures 5-28 and 5-27, we can see that the number of active participants is small and the level of turn-taking is low, compared to other conversation intervals. This might indicate that the participants were summarizing the event at this point instead of discussing new information.

Participant #6 stressed that, for her, the Hackathon was a success. The team created a mock-up that they can now discuss with the partners more easily, and

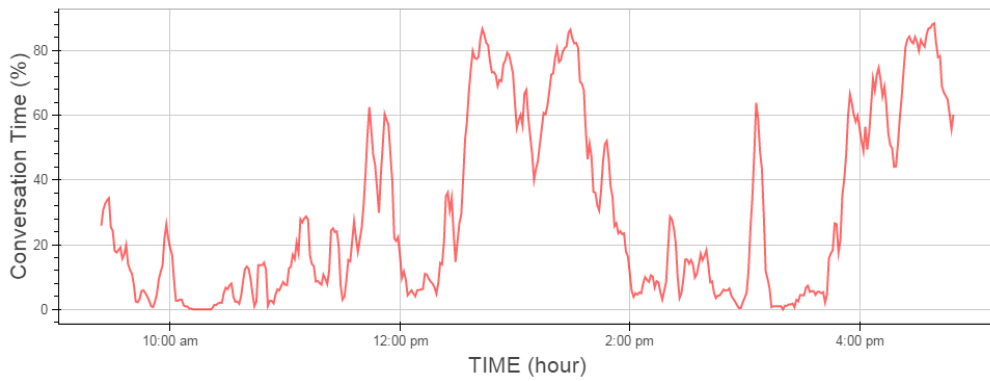


Figure 5-26: Percentage of conversation time per minute. Corporate Hackathon, Day 2, 9:20 am - 4:45 pm

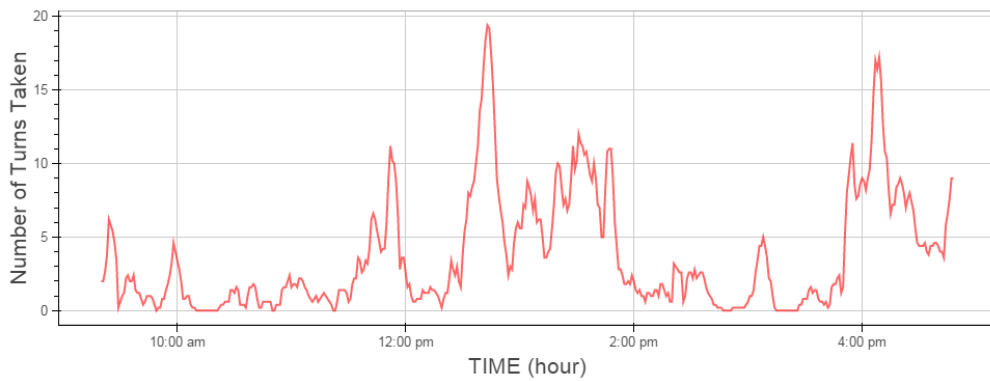


Figure 5-27: Number of turns taken per minute. Corporate Hackathon, Day 2, 9:20 am - 4:45 pm

better understand the team members' contributions and the potential implementation problems. The Hackathon ended at 4:45 pm.

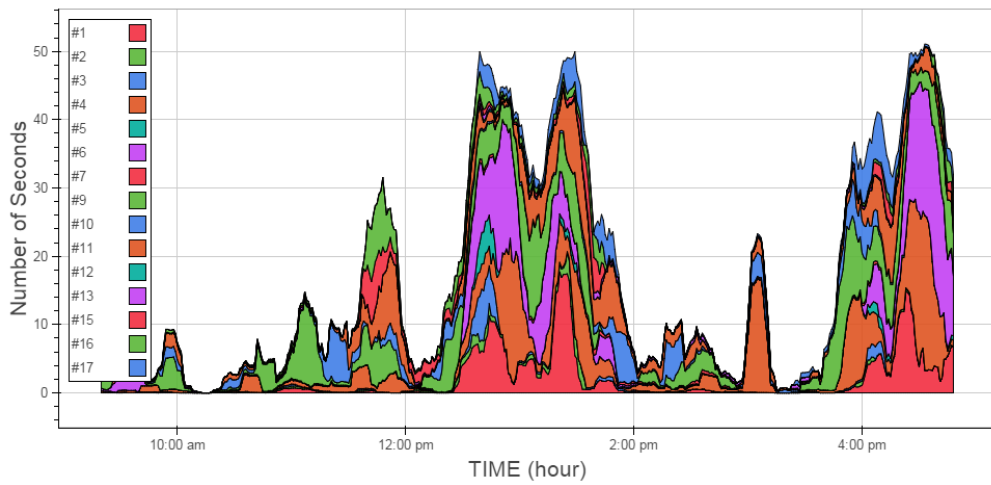


Figure 5-28: Contribution (in seconds) to the conversation, by participant. Corporate Hackathon, Day 2, 9:20 am - 4:45 pm



# Chapter 6

## Discussion

In our studies, we evaluated the proposed methodology for combining wearable technology with a new setting for studying team performance. We showed the richness of behaviors that can be observed in Hackathons and the potential for using our badges for future studies in this setting.

### 6.1 Badge Evaluation

The results of the studies in Chapter 5 show that our badges fulfill the main purpose for which they were created—they capture the main conversational patterns that we observed in the teams. Yet they have several limitations when used in the Hackathon setting. These limitations can be largely overcome by enhancing the badge design, as described in Chapter 7.

#### 6.1.1 Conversation Patterns

Using data on participants' vocalization, we measure participation, the parity of contribution to group discussions, and the level of turn taking. In some cases, the data can be used for measuring the frequency and duration of group meetings—a good proxy for the level of collaboration within the team.

### 6.1.2 Limitations

1. **Identifying sub-groups.** The main limitation of the badges is that they cannot be used for differentiating conversations that are held by sub-groups. Large teams often break into smaller working units which have their own dynamics and different efficacy, and therefore it is important to measure their behavior separately.
2. **Multiple Speakers.** Since we assume that only one person speaks at any given moment and identify that speaker by looking for the badge with highest volume, having multiple speakers registers as a high level of turn taking. Thus, arguments in which several people speak simultaneously, and similar behaviors, are not properly recorded by the badge.
3. **Base-Station Coverage.** The badges can only hold a limited number of samples in memory, roughly 10 minutes of data, and must therefore remain within close proximity to a base station. In a large-scale event, such as the medical Hackathon covered in this work, team members often moved outside the range of the base station, leading to data loss. This limitation has already been addressed by saving up to eight hours of data to the flash memory of the existing microcontroller.
4. **Other Behavioral Signals.** Other behaviors that we observed in the studies are not recorded by the badges. For example, participants' body language and eye contact may reveal level of interest in the conversation (e.g. who is listening to the discussion and who is checking email on a mobile device instead). Although these behaviors may be important for understanding group dynamics, they fall outside the scope of this work.

## 6.2 Hackathons as a Research Setting

Our observations show that the Hackathon is a rich setting for studying group dynamics in innovation teams, with opportunities to perform interventions and measure

the effect on team efficacy.

## 6.2.1 Observed Team and Teamwork Characteristics

### Working Modes

We identify several types of working modes in a Hackathon, namely group discussions, small talk, break-out sessions, and implementation time. These modes varied in the levels of different conversational attributes, such as the number of active speakers, the level of turn taking and how balanced the contribution of the participants was.

1. **Group discussions.** In a group discussion, the team members meet in order to share information and opinions, debate different options and decide on next steps. Because of the richness of vocal signals, this mode provides the researchers with the most data on group dynamics.

The teams that we observed showed two main types of group discussions and a similar pattern—a problem-definition and brainstorming meeting, followed by a series of coordination meetings.

The first meeting was typically longer than the other meetings, as team members had an opportunity to contribute their unique pieces of information to the conversation and discuss the best course of action. By introducing themselves and their expertise they built familiarity and trust that enable the participants to coordinate their next activities better [39, 73].

The literature on team performance also suggests that the behavior of the team during these initial meetings may be crucial for the success of future interactions, as they help in establishing a common language and understanding, and determine team dynamics [26, 36].

Status meetings were shorter, more focused, and varied in level of turn taking, participation and duration. In future studies we will use these attributes, and others, to investigate how the behavior of the team changes over time, and how it is affected by time pressure and critical meetings. As an example, we have

already observed higher levels of activity during coordination meetings that took place just before pitch practices.

2. **Small talk.** In these short sessions, small groups of participants engage in mostly off-topic, casual conversation. Small talk is an integral part of socializing, as it helps team members get to know each other better [26, 19, 69]. The previous chapter shows that these conversations often precede group meetings and are also common during breaks. Due to the limitations of the badges, we cannot differentiate between multiple conversations that take place at the same time, and therefore small talk registers as a period of high vocal activity with a high level of turn taking. Future versions of the badges with proximity detection will be able to differentiate between the simultaneous conversations.
3. **Breakout sessions.** A breakout session is a period of time in which the team breaks into small groups to discuss different aspects of the project. As we observed in both Hackathons, this behavior seems to help the team utilize its resources better by enabling it to work on multiple parts of the project at the same time. This mode is characterized by a more active conversation (compared to regular group discussions).
4. **Implementation.** When they were not having a discussion, the participants were typically busy working on the implementation of their ideas—either individually, as part of a smaller group, or together with the entire group. The amount of time dedicated to implementation varied across time and teams.

The duration, frequency and order of these work modes, as well as the group dynamics within each session, are likely to affect the team performance [86, 69].

### **Coordination and Sub-Groups**

Some of the groups showed a tendency to break into sub-groups at different stages of the Hackathons. This observation raises questions about the mechanisms that lead to this behavior, as well as the role of coordination in determining group efficacy—How



likely are teams of different sizes to break? What causes large teams to break into smaller working units? Does it improve their productivity, and if so, how can we encourage it?

The team in the corporate Hackathon was relatively large, up to 15 participants in the group discussions, and split into smaller working units several times. On one occasion, the participants were randomly assigned to teams as part of a breakout session planned by the organizers of the event. Later, after the main group discussions, some participants stayed to work on implementation while others went back to their offices and occasionally returned for a status meeting.

When we discussed the average size of teams with the organizers of the medical Hackathon, we were told that the average team has five to six members and that larger teams often struggle or break into completely separate teams. Nevertheless, the second team in that study had nine active members and seemed to perform well. It appears that the team realized in an early stage that in order to leverage its size, it must divide the work among several independent working units in different locations. This finding is supported by the literature on team size [27, 40, 48]. While this decision probably enabled each working unit to focus on its own tasks and work more efficiently, the data for this team reveal a low meeting frequency, a finding that might indicate a lack of coordination.

Occasional coordination meetings are essential for the success of complex tasks [47, 86]. Various collaboration tools, such as shared code repositories, e-mail, and conference calls make it easier for the participants to communicate and share their work, but these tools are not as powerful as face-to-face meetings and do not replace oral communication [69, 19, 47]. Our observations confirm that interpersonal communication varies throughout the event and among different teams.

### **Team Composition and Group Size**

While large teams enjoy a potentially larger pool of knowledge, smaller teams require less coordination and can therefore make decisions more easily. Our observations already showed a faster working pace in small groups. Using data from the naturally

occurring Hackathon teams of different sizes, future studies can determine whether there is a correlation between the number of participants and the success of their team.

Hackathon teams also showed different ranges of skills and backgrounds, and different levels of familiarity among team members. Further studies are required to validate this observation using a larger sample size, and to answer some of the questions about the relationship between team composition and performance: Is it really better to have a diverse team in a Hackathon? Can a group of friends work better than strangers? Are they just as likely to come up with a good idea? How should teams balance these two conflicting forces—diversity of skills and homogeneity of personality and background?

The Marshmallow Challenge may be an alternative environment for testing some of these effects. Due to the even shorter duration and the similarity to the planning and execution phases in Hackathons, Marshmallow Challenge teams can be seen as the fruit-fly version of Hackathon teams. These characteristics enable us to explore these questions in an environment where we can control the composition of the teams.

## **6.2.2 Potential Interventions**

### **Ice-Breaking Exercises**

Ice-breakers, such as the pairing exercise from the corporate Hackathon, can potentially improve team formation and quickly create stronger personal ties. By improving our badges to detect proximity to other participants, we will be able to tell whether new ties were created and how long the effect of the intervention lasted.

The method in which pairs are selected is expected to influence the effectiveness of the intervention. In the example from the corporate Hackathon, instead of randomly picking a partner, most people paired with the person sitting next to them. While this method should create new social connections, one can only assume that some of them would have been created even without an intervention, since people sitting next to each other are likely to talk. Creating truly random pairs should increase the

number of potential new social ties.

This intervention may be used to affect the composition of teams. If, for instance, future studies show that these pairs are more likely to form teams in Hackathons, then organizers will be able to create more diverse teams by randomly pairing people with different skills and backgrounds.

### **Energizing Activities**

In the medical Hackathon, the dancing activity is an example of an event-wide intervention intended to wake up and energize the teams. Yet, as the data from our badges suggest, it interrupted the team discussion and did not necessarily lead to more active behavior. By collecting behavioral data from more teams, researchers will be able to measure effect of such interventions and determine whether they are effective.

### **Interaction With Mentors**

Interactions with Hackathon mentors, such as pitch practices and status meetings, can be seen as potential interventions. They can be used to alter the structure of the event, and their timing may change the frequency with which the team members meet and coordinate their efforts.

## **6.2.3 Recommendations for Future Studies in Hackathons**

Based on our experience, we recommend making the study part of the official Hackathon agenda, and asking the organizers to introduce the researchers during the kick-off presentations. This will ensure that the Hackathon participants are fully aware of the research and make it easier to recruit them for the study.

We also recommend asking participants to opt in to the experiment and instrumenting them with badges before teams are formed. This will improve data quality, ensure data availability for early team discussions, and obviate the need for interruptions during these critical stages.



# Chapter 7

## Conclusions

In this thesis, we evaluated a methodology that combines an affordable wearable technology and the Hackathon setting to create a new environment for studying team performance. This combination reduces the time and burden involved in such studies and offers new opportunities for examining the effects of interventions on teamwork.

We presented a design for a wearable device that gathers vocalization data and used it to analyze the conversation dynamics of teams at several different events. Its simplicity and low price will enable future researchers to create their own badges and use them for similar studies.

Observations from our Hackathon studies found them to be potential settings for learning how team members interact and coordinate their efforts in order to achieve a shared goal. Using the badges, we measured different aspects of team behavior and showed how they change and evolve throughout the event.

To complete the evaluation of this methodology, we are planning to continue this work with a full deployment of the system in a Hackathon with hundreds of participants. All participants who choose to be part of this study will be instrumented with our new, affordable badges, allowing us to collect behavioral data from dozens of team simultaneously

## 7.1 Future Work

The following sections list some of the elements omitted due to time constraints and the limited scope of the project, and include our suggestions for future work in the field.

### Badge Design

- Eliminate the requirement for base-station proximity. We have already started experimenting with a new version of the badge that utilizes the existing flash memory on the microcontroller and saves up to eight hours of data
- Modify the badge so it can detect other badges and measure their signal strength. This will enable us to identify moments when the team splits into smaller working units
- Similarly, by measuring the signal strength of surrounding iBeacons (Bluetooth broadcasting units), we will be able to triangulate the location of the badges and study how people interact with space
- Improve the platform so it can better handle simultaneous speakers. One possible approach is to monitor the activity on multiple frequency bands
- Further reduce the price by replacing the BLE/microcontroller module and replacing the microphone breakout board with a digital microphone
- Add an accelerometer to the badge so we can sense movement and excitement

### Hackathons

- Large scale deployments of our system are required to measure the performance of multiple teams in the same event, compare the behaviors of these groups and learn how they correlate with the level of success
- Recruiting participants and instrumenting them with badges at an early stage of a Hackathon can reveal new information on team formation—how many people

do participants talk to before they decide which team to join? How long does it take them to decide? Can we create interventions that affect this process? Does the physical layout of the event affect it?

- Hackathons themselves can be studied as a new setting for innovative work. Researchers have recently begun to unlock the black box containing the dynamics of social network formation within corporations, showing that the mechanisms generating ties are frequently orthogonal to and sometimes at odds with the goals of the organization[50]. Corporate Hackathons provide a low-cost mechanism for organizations to potentially "rewire" their networks and for scholars to observe the capacity for organizations and individuals to strategically manipulate network formation[78]. Using our methodology, researchers can measure the effectiveness of these events and improve them
- Instrumenting mentors and judges with our badges during judging sessions can provide new information on the decision-making process of a group of experts asked to assess the creativity of an idea and its likelihood of success. This process is similar to those that take place in other innovative settings, such as venture capital firms, accelerators and incubators.

### **The Marshmallow Challenge**

The Marshmallow Challenge can be used for studying the effects of team composition and size on overall performance. We suggest a larger, more elaborate experiment in which participants will be randomly assigned into teams based on their background and demographics, thus creating teams of varying sizes and compositions.

It is also possible to use this activity to test the effect that teams have on each other. For example, we can create a sense of rivalry by having multiple teams work on the challenge simultaneously.





# Appendix A

## Schematic Diagrams and PCB Layout

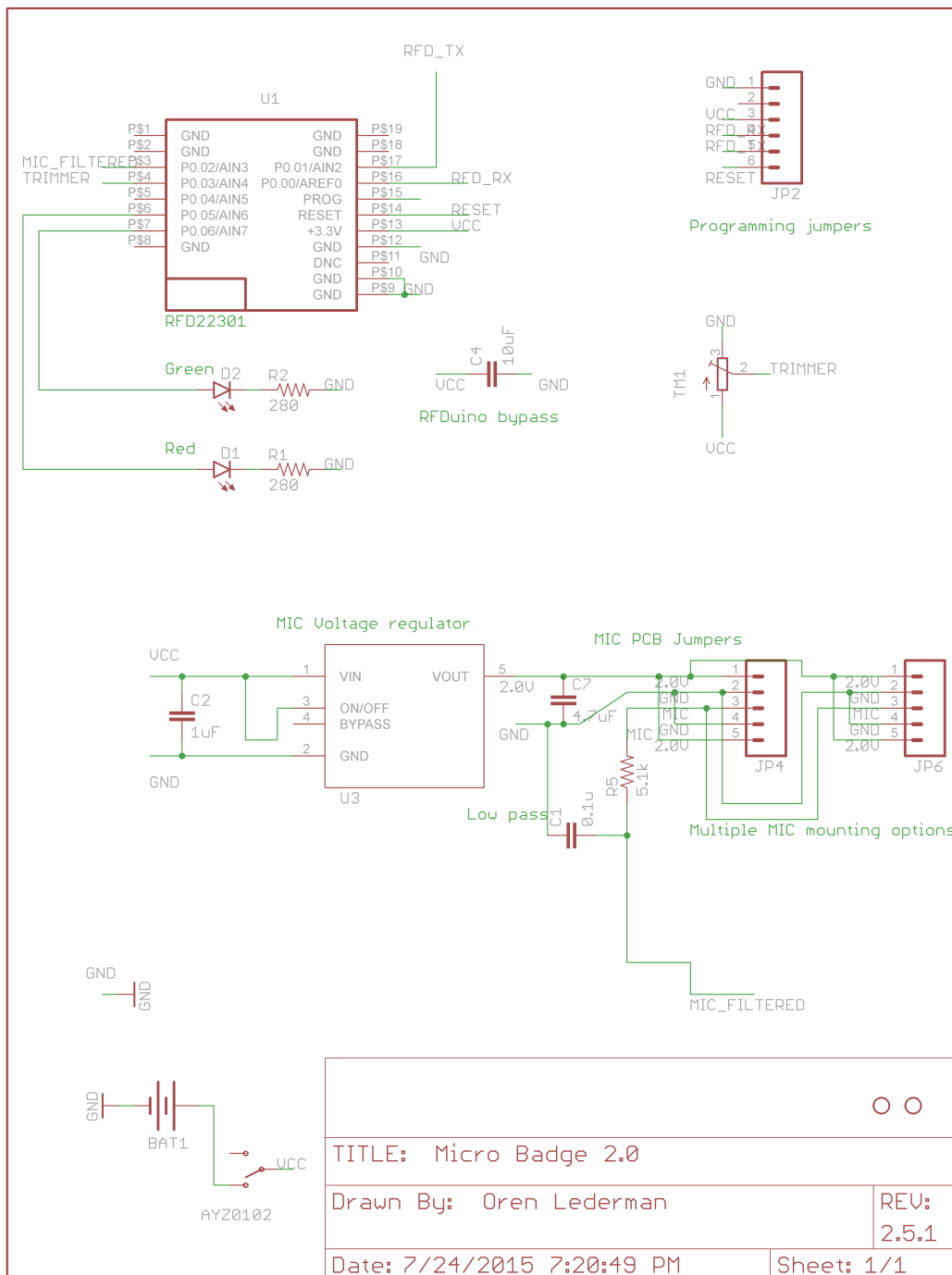


Figure A-1: Badge PCB schematic

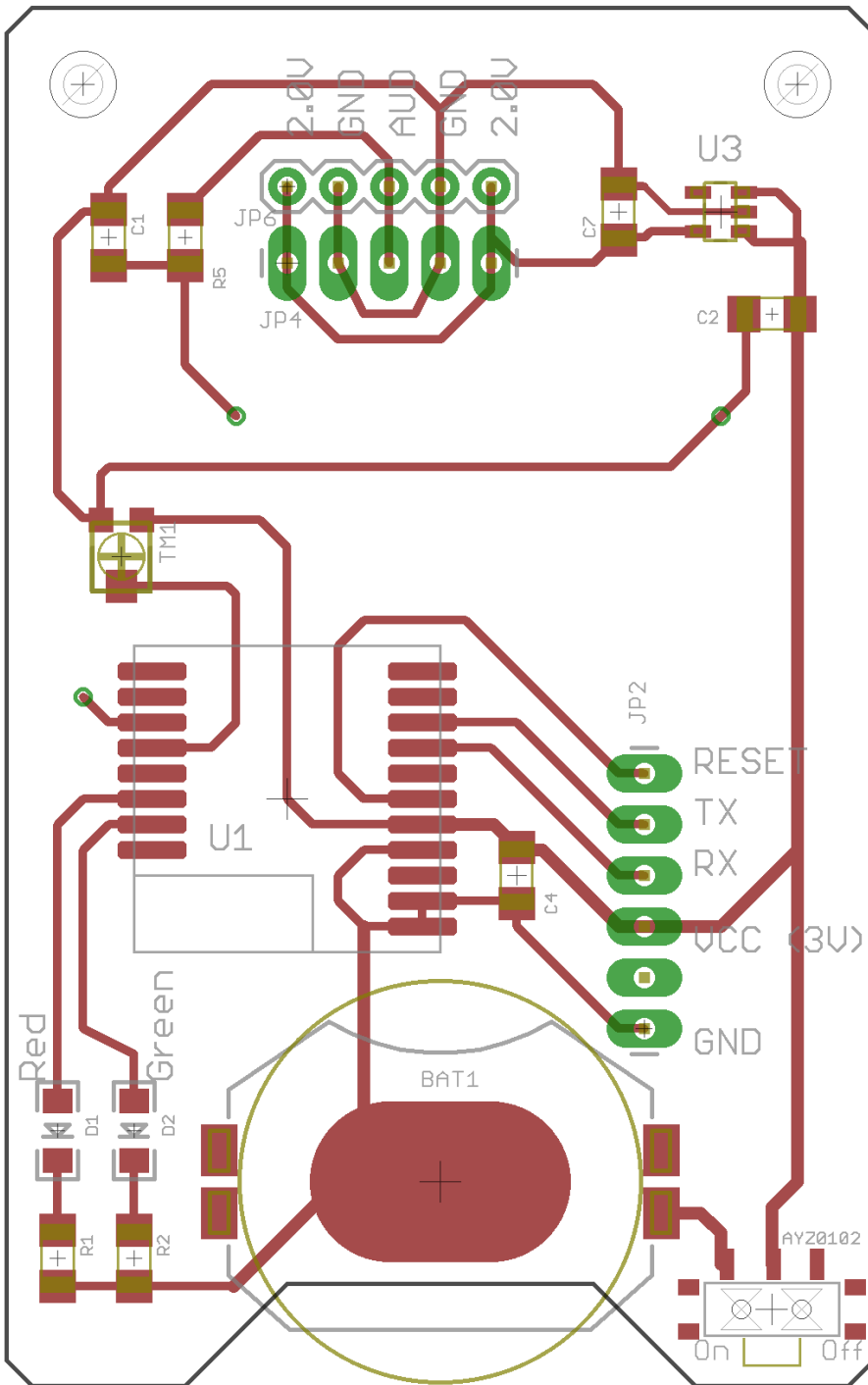


Figure A-2: Badge PCB top layer layout

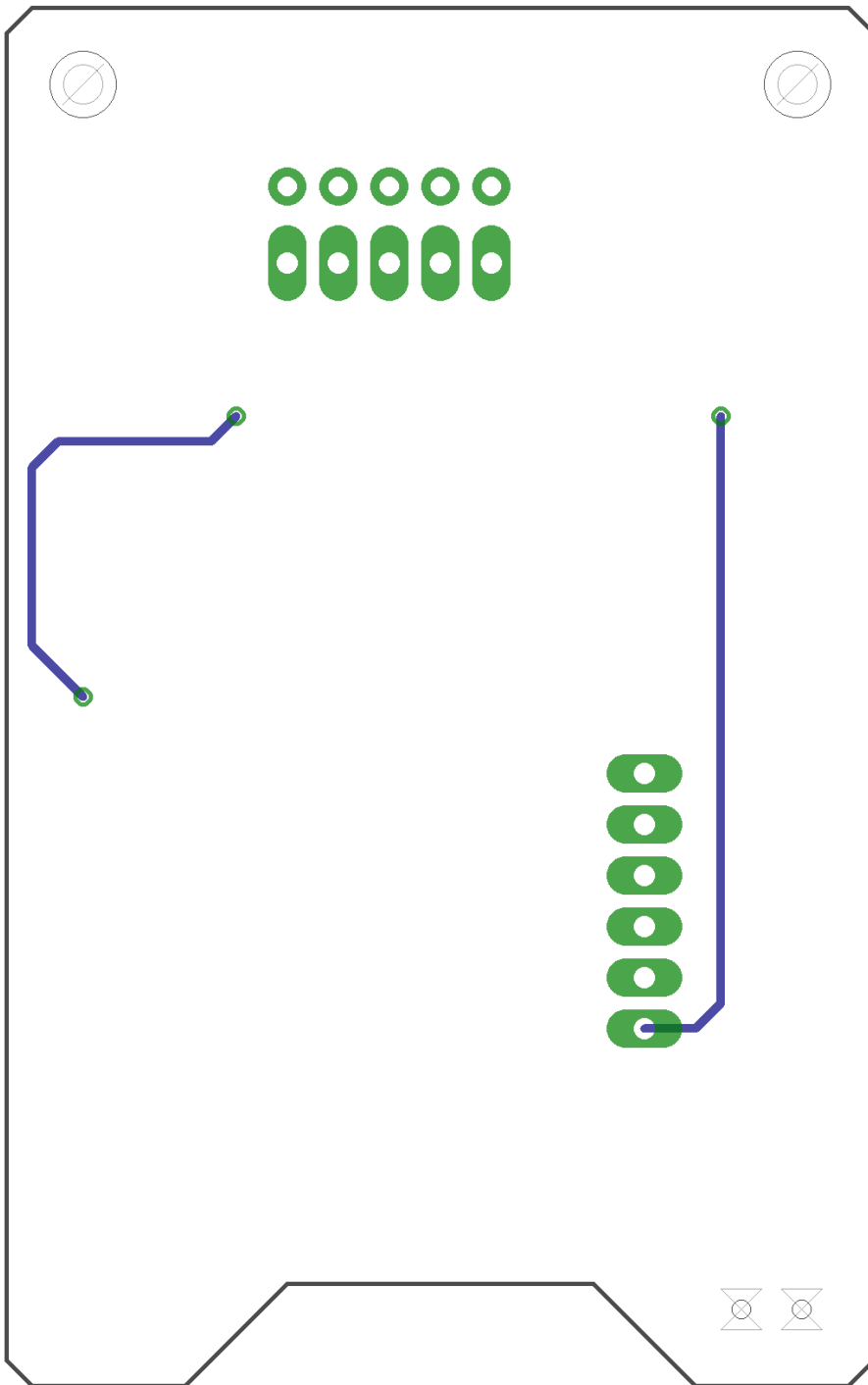


Figure A-3: Badge PCB bottom layer layout

# Appendix B

## Bill of Materials

| Parts    | Manufacturer                           | Manufacturer part # | Description   | Designator | Quantity per badge | Vendor   | Vendor part #        | Price |
|----------|--|---------------------|---|------------|--------------------|----------|----------------------|-------|
| 1        | RF Digital Wireless                    | RFD22301            | RFDuino   | U1         | 1                  | Mouser   | 975-RFD22301         | 14.99 |
| 2        | C&K Components                         | AZ70102AGRLC        | Power switch  | AZ70102    | 1                  | Mouser   | 611-AZ70102AGRLC     | 0.85  |
| 3        | Texas Instruments                      | LP2985AIM5-2.0/NOPB | LDO Voltage regulator                                 | U3         | 1                  | Mouser   | 926-LP2985AIM5Z0NOPB | 0.88  |
| 4        | Bourns                                 | 652-TC33X-2-103E    | Trimmer   | TM1        | 1                  | Mouser   | 652-TC33X-2-103E     | 0.27  |
| 5        |  | PPPC061LGBN-RC      | CONN FEMALE 6POS .100" R/A GOLD                       | JP2        | 1                  | DigiKey  | 55481-ND             | 0.83  |
| 6        | Yageo                                  | RC1206FR-07280RL    | RES SMD 280 OHM 1% 1/4W 1206                          | R1, R2     | 2                  | DigiKey  | 311-280FRCT-ND       | 0.1   |
| 7        | Taiyo Yuden                            | GMK316F106ZL-T      | CAP CER 10UF 35V Y5V 1206                             | C4         | 1                  | DigiKey  | 587-1352-1-ND        | 0.28  |
| 8        | Taiyo Yuden                            | UMK316AB7475KL-T    | CAP CER 4.7UF 50V 10% X7R 1206                        | C7         | 1                  | DigiKey  | 587-2994-1-ND        | 0.28  |
| 9        | TDK Corporation                        | C3216X7R1H105K160AB | CAP CER 1UF 50V 10% X7R 1206                          | C2         | 1                  | DigiKey  | 445-1423-1-ND        | 0.25  |
| 10       | Kemet                                  | C1206C104KARACTU    | CAP CER 0.1UF 250V 10% X7R 1206                       | C1         | 1                  | DigiKey  | 399-4674-1-ND        | 0.34  |
| 11       | Yageo                                  | RC1206FR-075K1L     | RES SMD 5.1K OHM 1% 1/4W 1206                         | R5         | 1                  | DigiKey  | 311-5.10KFRCT-ND     | 0.1   |
| 12       | Lite-On Inc                            | LTST-C150KFKT       | LED YELLOW ORANGE CLEAR 1206 SMD                      | D1         | 1                  | DigiKey  | 160-1403-1-ND        | 0.37  |
| 13       | Lite-On Inc                            | LTST-C150GKT        | LED GREEN CLEAR 1206 SMD                              | D2         | 1                  | DigiKey  | 160-1169-1-ND        | 0.37  |
| 14       | MPD                                    | BK-912              | HOLDER COIN CELL 20MM SMD                             | BAT1       | 1                  | DigiKey  | BK-912-ND            | 0.33  |
| 15       | FCI                                    | 609-3461-ND         | BERGSTIK II .100" SR STRAIGHT                         | JP4        | 1                  | DigiKey  | 609-3461-ND          | 0.26  |
|          |  |                     |   |            |                    |          | price per unit       | 20.5  |
| Misc     | Manufacturer                           | Manufacturer part # | Description   | Designator | Quantity per badge | Vendor   | Vendor part #        | Price |
| 16       | SparkFun                               | BOB-09868           | SparkFun MEMS Microphone Breakout - INMP401 (ADMP401) |            | 1                  | Sparkfun | BOB-09868            | 8.96  |
|          |  |                     |   |            |                    |          | price per unit       | 8.96  |
| PCB cost |  |                     |   |            |                    |          |                      |       |
| 17       | 10 PCBs, Red, 5x10cm, Lead free finish |                     |   |            |                    |          |                      |       |
|          | shipment,DHL                           |                     |   |            |                    |          |                      |       |
|          | Solder Paste Stencil                   |                     |   |            |                    |          |                      |       |
|          |  |                     |   |            |                    |          | Price per unit       | 7.775 |

Total price per unit \$37.24

Figure B-1: Bill of Materials

# Bibliography

- [1] ADMP401 Datasheet. <http://www.analog.com/media/en/technical-documentation/obsolete-data-sheets/ADMP401.pdf>.
- [2] Arduino homepage. <https://www.arduino.cc>.
- [3] BluePy Github repository. <http://github.com/IanHarvey/bluepy>.
- [4] Bluetooth Smart. <http://www.bluetooth.com/Pages/Bluetooth-Smart.aspx>.
- [5] BlueZ, the official linux bluetooth protocol stack. <http://www.bluez.org/>.
- [6] Hacking medicine grand hack 2015. <http://hackingmedicine.mit.edu/grandhack/>.
- [7] LP2985LV Datasheet. <http://www.ti.com/lit/ds/symlink/lp2985lv-n.pdf>.
- [8] Open beacon project. <http://www.openbeacon.org/>.
- [9] Raspberry Pi homepage. <http://www.raspberrypi.org>.
- [10] RFD22301 Datasheet. [http://www.rfduino.com/wp-content/uploads/2014/03/RFD22301.Data\\_Sheet\\_.11.24.13\\_11.38PM.pdf](http://www.rfduino.com/wp-content/uploads/2014/03/RFD22301.Data_Sheet_.11.24.13_11.38PM.pdf).
- [11] Sociometric solutions. <http://www.sociometricsolutions.com/>.
- [12] SparkFun MEMS microphone breakout - INMP401 (ADMP401). <http://www.sparkfun.com/products/9868>.
- [13] Nadav Aharony. *Social fMRI: measuring and designing social mechanisms using mobile phones*. PhD thesis, Massachusetts Institute of Technology, 2012.
- [14] Deborah Gladstein Ancona and David F Caldwell. Demography and design: Predictors of new product team performance. *Organization science*, 3(3):321–341, 1992.
- [15] Kenneth J. Arrow. The limits of organization. 1974.
- [16] Ronald J Baken and Robert F Orlikoff. *Clinical measurement of speech and voice. 2nd edition*. Singular Thomson Learning; San Diego, 2000.

- [17] Alain Barrat, Ciro Cattuto, Vittoria Colizza, Jean-Francois Pinton, Wouter Van den Broeck, and Alessandro Vespignani. High resolution dynamical mapping of social interactions with active rfid. *arXiv preprint arXiv:0811.4170*, 2008.
- [18] Allen C Bluedorn, Daniel B Turban, and Mary Sue Love. The effects of stand-up and sit-down meeting formats on meeting outcomes. *Journal of Applied Psychology*, 84(2):277, 1999.
- [19] Cheryl A Bolstad and Mica R Endsley. Tools for supporting team collaboration. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 47, pages 374–378. SAGE Publications, 2003.
- [20] Gerard Briscoe and Catherine Mulligan. Digital innovation: The hackathon phenomenon. Technical report, CreativeWorks London Working Paper, 2014.
- [21] Salas Eduardo Cannon-Bowers, Janis A and Sharolyn Converse. Shared mental models in expert team decision making. *Individual and group decision making: Current issues*, 221, 1993.
- [22] Leo Anthony Celi, Andrea Ippolito, Robert A Montgomery, Christopher Moses, and David J Stone. Crowdsourcing knowledge discovery and innovations in medicine. *Journal of medical Internet research*, 16(9), 2014.
- [23] Kenneth C Cheung, Stephen S Intille, and Kent Larson. An inexpensive bluetooth-based indoor positioning hack. In *Proceedings of UbiComp*, volume 6. Citeseer, 2006.
- [24] Tanzeem Choudhury and Alex Pentland. The sociometer: A wearable device for understanding human networks. In *CSCW'02 Workshop: Ad hoc Communications and Collaboration in Ubiquitous Computing Environments*, 2002.
- [25] Tanzeem Choudhury and Alex Pentland. Characterizing social interactions using the sociometer. In *Proceedings of NAACOS*. Citeseer, 2004.
- [26] Herbert H Clark and Susan E Brennan. Grounding in communication. *Perspectives on socially shared cognition*, 13(1991):127–149, 1991.
- [27] Luis A Cural, Rosalind H Forrester, Jeremy F Dawson, and Michael A West. It's what you do and the way that you do it: Team task, team size, and innovation-related group processes. *European Journal of Work and Organizational Psychology*, 10(2):187–204, 2001.
- [28] Yves-Alexandre de Montjoye, Arkadiusz Stopczynski, Erez Shmueli, Alex Pentland, and Sune Lehmann. The strength of the strongest ties in collaborative problem solving. *Scientific reports*, 4, 2014.
- [29] Artem Dementyev, Steve Hodges, Stuart Taylor, and Joshua Smith. Power consumption analysis of Bluetooth Low Energy, ZigBee and ANT sensor nodes in a cyclic sleep scenario. In *Wireless Symposium (IWS), 2013 IEEE International*, pages 1–4. IEEE, 2013.



- [30] Jacqueline W DePasse, Ryan Carroll, Andrea Ippolito, Allison Yost, Zen Chu, and Kristian R Olson. Less noise, more hacking: how to deploy principles from mit’s hacking medicine to accelerate health care. *International journal of technology assessment in health care*, 30(03):260–264, 2014.
- [31] Michael Diehl and Wolfgang Stroebe. Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of personality and social psychology*, 53(3):497, 1987.
- [32] Bryan D Edwards, Eric Anthony Day, Winfred Arthur Jr, and Suzanne T Bell. Relationships among team ability composition, team mental models, and team performance. *Journal of Applied Psychology*, 91(3):727, 2006.
- [33] Julie Fournet and Alain Barrat. Contact patterns among high school students. *PloS one*, 9(9):e107878, 2014.
- [34] H Clayton Foushee. Dyads and triads at 35,000 feet: Factors affecting group process and aircrew performance. *American Psychologist*, 39(8):885, 1984.
- [35] Connie JG Gersick. Time and transition in work teams: Toward a new model of group development. *Academy of Management journal*, 31(1):9–41, 1988.
- [36] Connie JG Gersick. Marking time: Predictable transitions in task groups. *Academy of Management journal*, 32(2):274–309, 1989.
- [37] Samuel D Gosling, Peter J Rentfrow, and William B Swann. A very brief measure of the big-five personality domains. *Journal of Research in personality*, 37(6):504–528, 2003.
- [38] Mark S Granovetter. The strength of weak ties. *American journal of sociology*, pages 1360–1380, 1973.
- [39] Deborah H Gruenfeld, Elizabeth A Mannix, Katherine Y Williams, and Margaret A Neale. Group composition and decision making: How member familiarity and information distribution affect process and performance. *Organizational behavior and human decision processes*, 67(1):1–15, 1996.
- [40] Roger Guimera, Brian Uzzi, Jarrett Spiro, and Luis A Nunes Amaral. Team assembly mechanisms determine collaboration network structure and team performance. *Science*, 308(5722):697–702, 2005.
- [41] Martine R Haas. Knowledge gathering, team capabilities, and project performance in challenging work environments. *Management Science*, 52(8):1170–1184, 2006.
- [42] Martine R Haas. The double-edged swords of autonomy and external knowledge: Analyzing team effectiveness in a multinational organization. *Academy of Management Journal*, 53(5):989–1008, 2010.

- [43] Josef Hallberg, Marcus Nilsson, and Kåre Synnes. Positioning with bluetooth. In *Telecommunications, 2003. ICT 2003. 10th International Conference on*, volume 2, pages 954–958. IEEE, 2003.
- [44] Randy Y Hirokawa. Discussion procedures and decision-making performance. *Human Communication Research*, 12(2):203–224, 1985.
- [45] L Richard Hoffman. Homogeneity of member personality and its effect on group problem-solving. *The Journal of Abnormal and Social Psychology*, 58(1):27, 1959.
- [46] Benjamin F Jones. Age and great invention. *The Review of Economics and Statistics*, 92(1):1–14, 2010.
- [47] Ralph Katz and Michael Tushman. Communication patterns, project performance, and task characteristics: An empirical evaluation and integration in an r&d setting. *Organizational behavior and human performance*, 23(2):139–162, 1979.
- [48] Jon R Katzenbach and Douglas K Smith. *The wisdom of teams: Creating the high-performance organization*. Harvard Business Press, 1993.
- [49] Bradley L Kirkman, Benson Rosen, Paul E Tesluk, and Cristina B Gibson. The impact of team empowerment on virtual team performance: The moderating role of face-to-face interaction. *Academy of Management Journal*, 47(2):175–192, 2004.
- [50] Adam M Kleinbaum, Toby E Stuart, and Michael L Tushman. Discretion within constraint: Homophily and structure in a formal organization. *Organization Science*, 24(5):1316–1336, 2013.
- [51] Jan Kratzer, Oger Th AJ Leenders, and Jo ML Van Engelen. Stimulating the potential: Creative performance and communication in innovation teams. *Creativity and Innovation Management*, 13(1):63–71, 2004.
- [52] Jan Kratzer, Roger Th AJ Leenders, and Jo ML Van Engelen. Team polarity and creative performance in innovation teams. *Creativity and Innovation Management*, 15(1):96–104, 2006.
- [53] Steven Leckart. The hackathon is on: Pitching and programming the next killer app. *Wired, San Francisco*, 17, 2012.
- [54] Steven Leckart. The hackathon fast track, from campus to silicon valley. <http://www.nytimes.com/2015/04/12/education/edlife/the-hackathon-fast-track-from-campus-to-silicon-valley.html>, 2015.
- [55] Jerad Lewis and Brian Moss. MEMS Microphones, the future for hearing aids. [http://www.analog.com/library/analogdialogue/archives/47-11/hearing\\_aids.pdf](http://www.analog.com/library/analogdialogue/archives/47-11/hearing_aids.pdf), 2013.

- [56] Beng-Chong Lim and Katherine J Klein. Team mental models and team performance: a field study of the effects of team mental model similarity and accuracy. *Journal of Organizational Behavior*, 27(4):403, 2006.
- [57] Eugene Mandel. How the napa earthquake affected bay area sleepers. <https://jawbone.com/blog/napa-earthquake-effect-on-sleep/>, 2014.
- [58] Nathan J McNeese, Madhu C Reddy, and Evan M Friedenberg. Towards a team mental model of collaborative information seeking during team decision-making. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 58, pages 335–339. SAGE Publications, 2014.
- [59] Xavier Anguera Miro, Simon Bozonnet, Nicholas Evans, Corinne Fredouille, Gerald Friedland, and Oriol Vinyals. Speaker diarization: A review of recent research. *Audio, Speech, and Language Processing, IEEE Transactions on*, 20(2):356–370, 2012.
- [60] Susan Mohammed, Lori Ferzandi, and Katherine Hamilton. Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management*, 2010.
- [61] Richard Newton. The hackathon enters the corporate mainstream. <http://www.ft.com/intl/cms/s/0/88212d8a-a716-11e4-8a71-00144feab7de.html#axzz3eZ7XESZL>, 2015.
- [62] Gerardo A Okhuysen and Kathleen M Eisenhardt. Integrating knowledge in groups: How formal interventions enable flexibility. *Organization Science*, 13(4):370–386, 2002.
- [63] Gerardo Andrés Okhuysen. Structuring change: Familiarity and formal interventions in problem-solving groups. *Academy of Management Journal*, 44(4):794–808, 2001.
- [64] Daniel Olguín Olguín. *Sociometric badges: wearable technology for measuring human behavior*. PhD thesis, Massachusetts Institute of Technology, 2007.
- [65] Daniel Olguín-Olguín and Alex Pentland. Sensor-based organisational design and engineering. *International Journal of Organisational Design and Engineering*, 1(1):69–97, 2010.
- [66] Charles A O’Reilly III, Katherine Y Williams, and Sigal Barsade. Group demography and innovation: Does diversity help? 1998.
- [67] Sung Park, Ivo Locher, Andreas Savvides, Mani B Srivastava, Alvin Chen, Richard Muntz, and Spencer Yuen. Design of a wearable sensor badge for smart kindergarten. In *Wearable Computers, 2002.(ISWC 2002). Proceedings. Sixth International Symposium on*, pages 231–238. IEEE, 2002.

- [68] Alex Pentland. Socially aware, computation and communication. *Computer*, 38(3):33–40, 2005.
- [69] Alex Pentland. The new science of building great teams. *Harvard Business Review*, 90(4):60–69, 2012.
- [70] Alex Pentland. *Social Physics: How Good Ideas Spread-The Lessons from a New Science*. Penguin, 2014.
- [71] Alex Pentland and Tracy Heibeck. *Honest signals*. MIT press Cambridge, MA, 2008.
- [72] Vivek Ravisankar. Hacking hackathons: How 4,000 events taught me the four keys for any organization to be a successful host. <http://www.forbes.com/sites/vivekravisankar/2015/03/04/hacking-hackathons-how-4000-events-taught-me-the-four-keys-for-any-organization-2015>.
- [73] Ray Reagans, Linda Argote, and Daria Brooks. Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management science*, 51(6):869–881, 2005.
- [74] Lionel P Robert Jr, Alan R Dennis, and Manju K Ahuja. Social capital and knowledge integration in digitally enabled teams. *Information Systems Research*, 19(3):314–334, 2008.
- [75] Barry R Schlenker, Bob Helm, and James T Tedeschi. The effects of personality and situational variables on behavioral trust. *Journal of personality and social psychology*, 25(3):419, 1973.
- [76] Scott Shane. Prior knowledge and the discovery of entrepreneurial opportunities. *Organization science*, 11(4):448–469, 2000.
- [77] Phill Smith. Comparisons between low power wireless technologies. *US Patent CS-213199-AN*, 2011.
- [78] Toby E Stuart and Olav Sorenson. Strategic networks and entrepreneurial ventures. *Strategic Entrepreneurship Journal*, 1(3-4):211–227, 2007.
- [79] Martin Szomszor, Patty Kostkova, Ciro Cattuto, Wouter Van den Broeck, Alain Barrat, and Harith Alani. Providing enhanced social interaction services for industry exhibitors at large medical conferences. In *Developments in E-systems Engineering (DeSE), 2011*, pages 42–45. IEEE, 2011.
- [80] Ingo R Titze. Principles of voice production. 1994.
- [81] Sue E Tranter, Douglas Reynolds, et al. An overview of automatic speaker diarization systems. *Audio, Speech, and Language Processing, IEEE Transactions on*, 14(5):1557–1565, 2006.

- [82] Brian Uzzi, Satyam Mukherjee, Michael Stringer, and Ben Jones. Atypical combinations and scientific impact. *Science*, 342(6157):468–472, 2013.
- [83] Jo ML Van Engelen, Derk Jan Kiewiet, and Pieter Terlouw. Improving performance of product development teams through managing polarity. *International Studies of Management & Organization*, pages 46–63, 2001.
- [84] Nicolas Voirin, Cécile Payet, Alain Barrat, Ciro Cattuto, Nagham Khanafer, Corinne Régis, Byeul Kim, Brigitte Comte, Jean-Sébastien Casalegno, Bruno Lina, et al. Combining high-resolution contact data with virological data to investigate influenza transmission in a tertiary care hospital. *Infection Control & Hospital Epidemiology*, pages 1–7, 2015.
- [85] Warren E Watson, Kamalesh Kumar, and Larry K Michaelsen. Cultural diversity’s impact on interaction process and performance: Comparing homogeneous and diverse task groups. *Academy of management journal*, 36(3):590–602, 1993.
- [86] Suzanne Weisband. Maintaining awareness in distributed team collaboration: Implications for leadership and performance. *Distributed work*, pages 311–333, 2002.
- [87] Katherine Y Williams and Charles A O’Reilly. Demography and diversity in organizations: A review of 40 years of research. *Research in organizational behavior*, 20:77–140, 1998.
- [88] Katherine A Wilson, Eduardo Salas, Heather A Priest, and Dee Andrews. Errors in the heat of battle: Taking a closer look at shared cognition breakdowns through teamwork. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(2):243–256, 2007.
- [89] Anita Williams Woolley, Christopher F Chabris, Alex Pentland, Nada Hashmi, and Thomas W Malone. Evidence for a collective intelligence factor in the performance of human groups. *science*, 330(6004):686–688, 2010.
- [90] Stefan Wuchty, Benjamin F Jones, and Brian Uzzi. The increasing dominance of teams in production of knowledge. *Science*, 316(5827):1036–1039, 2007.
- [91] Tom Wujec. The marshmallow challenge. <http://marshmallowchallenge.com/>.
- [92] Tom Wujec. TED Talk: Build a tower, build a team. [http://www.ted.com/talks/tom\\_wujec\\_build\\_a\\_tower](http://www.ted.com/talks/tom_wujec_build_a_tower), 2010.