# Visiting the 'Lie-brary:' Exploring Data Engagement as Participant Incentive

**Madeline E. Smith, Asmaa Aljuhani, Jeremy Birnholtz**
Communication Studies Department
Northwestern University
Evanston, IL
madsesmith@u.northwestern.edu
asmaaaljuhani2012@u.northwestern.edu
jeremyb@northwestern.edu

**Jeff Hancock, Lindsay Reynolds**
Department of Communication
Department of Information Science
Cornell University
Ithaca, NY
jth34@cornell.edu
llr48@cornell.edu


*CSCW'14 Companion,* Feb 15-19 2014, Baltimore, MD, USA
ACM 978-1-4503-2541-7/14/02.
http://dx.doi.org/10.1145/2556420.2556786

## Abstract
As people's online behavior increasingly leaves traces behind, it is tempting for researchers to gather and analyze these traces. This raises both ethical and logistical challenges in gathering and storing data; and in motivating people to share their data. We report on our experience developing an Android OS app to gather text messages and information about deception, and a web visualization interface that allows participants to engage with their data. We discuss our experiences and challenges, and solicit feedback and design ideas.

## Author Keywords
Text messaging, mobile apps, CMC, data engagement

## ACM Classification Keywords
H.5.m. Information interfaces and presentation

## Introduction & Background
It has become clear in recent years that the digital traces left by online interactions allow for studies of human communication at unprecedented levels of detail. Where prior studies relied on people's recollections in interviews or ability to record interactions in a diary, it is now possible to record and analyze interactions at the message or utterance level. This enables unprecedented examination of behaviors and communication strategies.

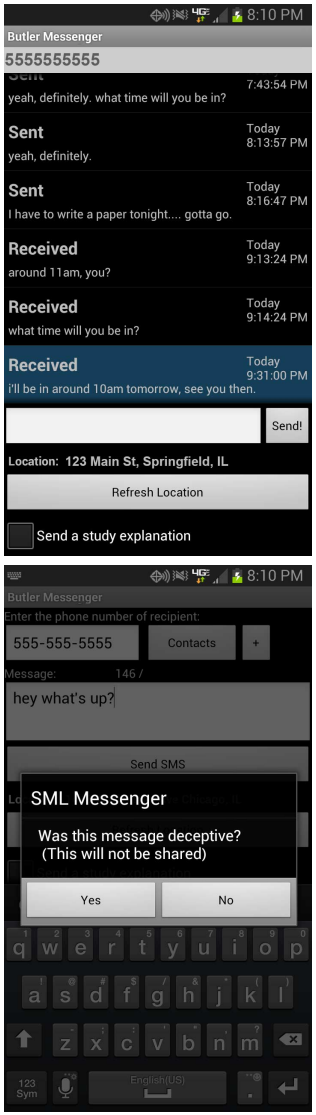


**Figure 1:** Screenshots of the Butler Messenger App

Even as logging all of people's digital interactions is a tempting solution to myriad research challenges, however, this method also raises both ethical and logistical challenges. There are ethical concerns around the storage of identifying information about non-consenting third parties, and capturing the intimate details of people's interactions. Logistically there are challenges in getting people to use a message logging application, in terms of incentives and recruitment.

Related to this, a recent trend in mobile applications is the capacity to engage with and reflect on traces of behavior. Popular applications such as Moves[1], Sleep Cycle[2], and Mint[3] capture data about physical movement, sleep, and financial transactions, and allow users to track, visualize, and engage with these data to better understand their behavior and habits. These and other personal informatics tools "help people collect personal information to improve self-knowledge" [1].

We wondered if we could exploit the confluence of these trends by designing an app that would capture message-level SMS text message data in a manner sensitive to ethical concerns, and use the capacity for engagement and visualization of one's data as an incentive for participation in our study. To do so we built an Android mobile application combined with a web survey for collecting text messages and information about deceptive behavior, and developed a web interface called the "Lie-brary" with which participants can engage with and visualize data about their messaging behavior.

## Our Method and Systems

*Butler Messenger*

To gather text messages and information about whether these messages were deceptive, we developed Butler Messenger (named for our "butler lies" project), an Android OS app. The app is functionally similar to the default Android text messaging application. It allows for sending and receiving text messages, accessing the contact list, and viewing messages in ongoing conversations (see Figure 1, top). There are a few key differences, however.

First, the app displays a pop-up after each message is sent that asks whether the message is deceptive or not (see Figure 1, bottom). Note that this information is not shared with message recipients or stored on the phone.

Second, the app sends all sent and received messages to our secure database. To avoid storing identifying information, sender and recipient phone numbers are hashed, and known proper names are automatically removed from messages. We also record the date, time, and GPS location of the phone.

After participants register for the study (see [2]) and complete a screening questionnaire, they are given instructions on how to download and install the app. They are then instructed to use the app as a substitute for their regular text messaging application for one week. Using cash payment as an incentive and with IRB approval from our respective institutions, we have run approximately 200 participants through this protocol, and gathered ~15,000 text messages.

[1] http://www.moves-app.com/

[2] http://www.sleepcycle.com/

[3] http://www.mint.com/

At the conclusion of one week, participants complete a web questionnaire. For each conversation they are asked about the nature (i.e., friend, family, colleague, etc.) and closeness of their relationship with the message recipient. For the 10 locations they most commonly sent messages from they are asked about the type of place that location is (using categories borrowed from Foursquare). And for messages indicated to be deceptive, they are asked what about the message is deceptive and how bad they would feel if the recipient found out the message is a lie.

*The Lie-brary*

To allow participants to engage with their data and hopefully provide incentive for future participants to enroll, we developed a web-based data engagement and visualization interface called the "Lie-brary." The site is implemented using PHP, MySQL, and JavaScript; and uses APIs such as Google Charts and Google Maps.

Lie-brary allows participants to view and reflect on their deception and messaging behavior. Importantly, participants may access the site only after they have used Butler Messenger for the full study period and have completed the post-study questionnaire. This prevents any influence of the reflection on their behavior during the study itself (though, of course, their behavior after using Lie-brary is also of interest).

To access Lie-brary, participants sign in with their phone number (matched using a hash function for security) and a unique ID number. After logging in, participants see a screen that displays some of their data in chart form (see Figure 2) along with a very brief text summary, a click-able arrow that takes them to another visualization, and two types of navigation tools: a set of simple questions at the bottom that they can click on to explore these aspects of their data (Who? What? Where? When?) and a chart/map "switch" that toggles between viewing the data for each question in these two formats.

While we do not describe all of the possible visualizations here, we give several examples to provide a general sense of the system. For example, clicking the "Who?" button (Figure 2) shows the percentage of sent messages to different types of relations (e.g., friends, colleagues, etc.) that were and were not marked as lies. Selecting the map view (Figure 3) shows the locations where messages to these people were sent. Clicking a map pin displays the text of the message sent from that location.

Clicking the "What?" button, in turn, displays Word Clouds constructed from the words in both truthful and deceptive text messages. The "When?" button displays a chart showing at what time and on what days deceptive messages were sent.

The goal in all of these is to allow participants to engage with and explore their data and behavior. We are in the initial phases of gathering feedback on this system and interface from participants and aim to iteratively refine the design based on their feedback.

*Challenges*

Our design and implementation have faced myriad challenges ranging from not being compatible with all Android phones to the increased use of IP-based (i.e., non-SMS) messaging apps such as WhatsApp, which are not captured by our system. By demonstrating this system and engaging with the CSCW community we

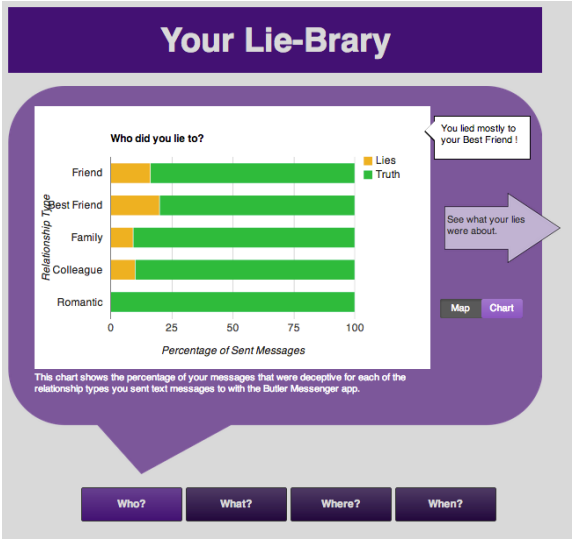


**Figure 2:** Chart showing message recipients.

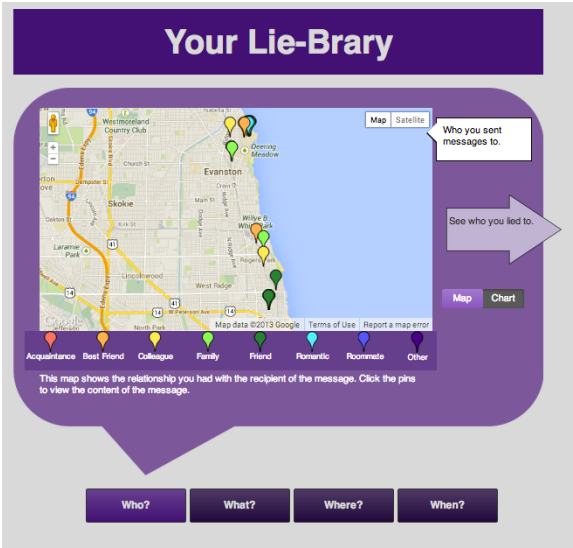


**Figure 3:** Map visualization of message recipients

hope to share our experience with these challenges to improve others' efforts as well as our own.

## Demo Goals and Conclusion

We have two goals in presenting this demo at CSCW. The first is to demonstrate and get feedback on a novel system and research method we have developed that we hope will allow for both gathering data and participants to reflect on their interaction behavior. The second goal is to allow CSCW attendees to directly experience this method and engagement with their own data during the conference itself. We invite those with Android phones to download and use Butler Messenger while at the conference, and use the Lie-Brary interface at the end to reflect on their experience and provide us with feedback. Those without Android phones can also engage with sample data at the demo reception.

## Acknowledgements

This work is supported by the National Science Foundation Grants IIS-0915081 and DGE-0824162. We thank Jessie Taft, Ashley Falcon, Yjaden Wood, and Scott Cambo for valuable assistance with this project.

## References

[1] Li, I., Dey, A., Forlizzi, J., Höök, K. and Medynskiy, Y. Personal informatics and HCI: design, theory, and social implications. In *Ext. Abstracts ACM CHI* (2011), 2417-2420.

[2] Smith, M. E., Birnholtz, J., Reynolds, L. and Hancock, J. People, place and time: the daily rhythms of deception in interpersonal text messaging. Paper presented at the Conference of the International Communication Association (2013).