

---

# Exploring Identity Signals on an Anonymous Mobile Posting App

## Joseph Seering

Human-Computer Interaction  
Institute  
Carnegie Mellon University  
5000 Ave.  
Pittsburgh, PA 15210 USA  
jseering@cs.cmu.edu

## Laura Dabbish

Human-Computer Interaction  
Institute  
Carnegie Mellon University  
5000 Ave.  
Pittsburgh, PA 15210 USA  
dabbish@cs.cmu.edu

## Abstract

Users are drawn to anonymous apps because of the appeal of the protection provided by anonymity or partial disclosure of identifying information. Partial or complete anonymity supports identity exploration and social boundary management but presents challenges in discerning user sincerity and expertise through conventional signals. This project focuses on understanding how users evaluate the posts of others in an anonymous and ephemeral setting on the mobile anonymous mobile posting app Yik Yak.

## Author Keywords

Anonymity, privacy, identity, ephemeral communication

## Introduction

A majority of internet users believe that people should have the ability to use the internet anonymously for certain activities [4]. Most internet users have taken steps to conceal their identity and personal information online, from clearing cookies to encrypting their email to using pseudonyms [6].

Research has shown users seek anonymity in many ways for a variety of reasons. Anonymity facilitates a variety of prosocial behaviors, including seeking and receiving advice in sensitive situations [2] and trying out different identities

[7]. However, anonymity can also facilitate harassment, bullying, and deception [1][3].

Yik Yak is an app that allows users to post short comments seen by other users of the app within a small geographic radius. This geographic location and the content of posts are the primary signals available to users attempting to interpret other users' sincerity and expertise. Yik Yak has been presented by some as representative of the dangers of anonymity online [3], with authors highlighting instances of harassment that could have been encouraged by the protection afforded by anonymity within the app.

This project explores how users interpret posts with very limited signals and how they rate posts based on their perception of the users intent. Understanding how users perceive and rate others posts will help inform the design of tools that allow users to clarify and validate signals about intent without sacrificing the protection that anonymity provides. Better informing users about posters intent will allow them to address some of the negative behaviors within their communities by moderating away insincere or disruptive content.

### Signals and Platform characteristics

Donath [1] explores the possibility of category deception, or pretending to be something that one is not, in online settings. She describes the difference between assessment and conventional signals, and explains how each piece of information available from Usenet posts might fit into each category. Many online signals, especially in discussion forums, are conventional signals, as in the case of a poster claiming to be an expert in a domain. Some platforms have developed modes for converting conventional signals into higher effort but more powerful

assessment signals, such as the attachment of a formal, proven tag of expertise [5]. On Yik Yak, the only types of identity knowledge that are available to users are location and post text.

The post in Figure 1 was observed at the center of the campus of a large public research university, meaning that the user who made the post must have been within a radius of a few miles of this location. No further identifying information, such as a username, is provided within the post itself beyond the text.

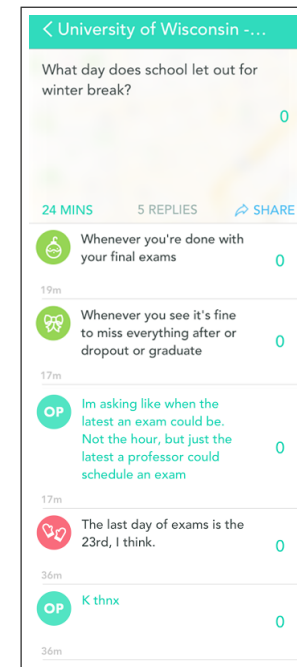


Figure 1: Sample Yik Yak post with replies

The comments below the post provide some additional

context; here, the user clarifies their question in the third reply. Users have persistent usernames in the comments in the form of randomly assigned pictorial representations. The original poster (OP) is marked with the OP tag and blue-green text.

In this example, the OP seeks information about the end of the semester, possibly for the purpose of arranging travel plans. Two respondents reply, addressing the original question in different ways. The OP then clarifies the question and receives a third reply. The OP implies acceptance of this third reply as factual.

The above post is one relatively low-stakes example of information seeking and sharing. If the OP were uncertain about the information provided, he or she could probably have looked online for a schedule of exams. Some posts on Yik Yak have much higher stakes, as in the case of a post seeking help or comfort after a traumatic event, e.g. Found out a friend of mine is considering going out with a guy who repeatedly raped me. I don't know what to do. An extreme example came during recent racially-charged events at the University of Missouri, where an anonymous poster wrote on Yik Yak, Im going to stand my ground tomorrow and shoot every black person I see. Assessing the intent of posters on Yik Yak and other anonymous communication apps is crucial to forming an appropriate response.

## Methods

We are studying how conventional signals develop in an anonymous communication setting like Yik Yak. In this study we are focusing on two primary research questions:

1. How is the participation dynamic of an anonymous communication community associated with

characteristics and composition of the local community?

- (a) How is the volume of posts related to type and size of local institutions?
  - (b) How do types of posts vary with community size? Homogeneity?
  - (c) What types of posts gather the most positive and negative feedback, and does this vary across different types of local institutions?
2. How do users interpret different types of posts?
    - (a) What features of posts do users look for to determine whether a post is sincere and/or informed?
    - (b) How do users infer identity and other characteristics of posters using information in the application environment?
    - (c) What are conventional versus assessment signals in anonymous communication applications and how do they develop?
    - (d) How do users structure their posts so as to achieve the desired response?

In this study, we are exploring posting norms across a variety of geographic locations on Yik Yak. We are initially focusing on a sample of four postsecondary institutions with substantially different characteristics. We have manually transcribed a large set of Yik Yak posts from these four institutions.

First, we are coding collected posts by topic in order to validate an existing coding scheme developed for posts on Whisper and Secret [2]. We are validating this coding

scheme by comparing codes assigned by two researchers to calculate a Cohen's Kappa. We will use this coding scheme to predict community response to the post in terms of score (total upvotes minus total downvotes).

In the second phase of the study we are interviewing Yik Yak users from each of these geographic locations. We will ask participants about how they perceive a variety of posts. We will ask them to describe what image they have of the person who made the post, what cues in the post led them to this image, and how sincere they find the post. We will develop a separate coding scheme for categorizing different types of conventional and assessment signals.

The end result of this project will be a validated coding scheme for categorizing different types of Yik Yak posts that has predictive power as a factor in a model of total score of the posts, as well as framework for understanding how users interpret identity and sincerity of posters in an environment where the only identity signals are geographic location and post text. These results will apply more broadly to the study of perception of anonymous posts online, and may inform the development of platforms incorporating new types of signals that clarify intent without sacrificing the protection afforded by anonymity.

## Conclusion

Understanding anonymous behavior online is more important than ever. Online harassment, cyberbullying, doxxing, swatting, and a plethora of other horrifying behaviors are constantly making news. At the same time, anonymity has provided tremendous value to marginalized groups and individuals seeking support. This project aims to contribute to an understanding of the ways in which users parse ephemeral posts from users with no persistent

identities. This knowledge will help design systems that facilitate positive interactions while removing ambiguity about the intent of users making misleading and harmful posts.

## References

- [1] Donath, J. *Identity and deception in the virtual community*. Routledge, London, 1998.
- [2] Kang, D., and Sutton. Why people use anonymous communication applications. In *Proc. CSCW*, ACM Press (2016).
- [3] Kowalski, and Limber. Psychological, physical, and academic correlates of cyberbullying and traditional bullying. *Journal of Adolescent Health* (2013).
- [4] Madden, M., and Rainie, L. Americans attitudes about privacy, security and surveillance. *PEW Research Center* (2015).
- [5] Morris, K. Reddit gets a bit of flair. *The Daily Dot* (2011). Retrieved from: <http://www.dailydot.com/news/reddit-gets-bit-flair/>.
- [6] Rainie, L. e. a. Anonymity, privacy, and security online. *PEW Research Center* (2013).
- [7] Yee, N. *The Proteus Paradox: How Online Games and Virtual Worlds Change Us-and how They Don't*. Yale University Press, 2014.