

Impact of Anonymity and Partial Anonymity on Social Interaction

A Research Paper submitted to The Department of
Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science.

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

Michael Yates

Spring 2022

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed, Michael Yates

Advisor

Sean M. Ferguson, Department of Engineering and Society

Impact of Anonymity and Partial Anonymity on Social Interaction

There are many potential benefits of allowing for complete anonymity in the facilitation of conversations or debates. Without fear of judgment, people are more willing to express their honest opinions. It also allows users to engage in self-discovery by asking the questions they wouldn't normally be comfortable saying aloud. On the contrary, it enables the spread of misinformation without fear of consequences and enables online trolling. Developers in modern times must determine to what extent a user's online identity will be made available to others on the platform.

Using Sociotechnical System's Politics of Design, we can analyze the design decision of anonymity, and its influence on people's behavior. Social media is unique in that the users take an active role in the production of the media, therefore directly impacting other users' online experience. To effectively guide users into healthy interactions within a platform, we must examine how app design affects platform culture, which ultimately dictates user behavior and interaction. One way of utilizing the politics of design is to study examples of similar productions, and understand the tradeoffs made by their design decisions (Gordon Baxter et al., 2010 p.7). There are many case studies of anonymity and the promotion of free speech, and how it can inadvertently enable cyberbullying. By considering design decisions, we can look at design decisions across different social media platforms, and their impact on toxicity and general app culture. I then locate developers as part of the app culture sphere to better understand their influence.

A great point of analysis is on the commonly used anonymous message app YikYak, which was developed in 2013 and allowed for the posting of anonymous messages where those in a similar location to the poster could see the message. Whisper is another anonymous social media, which pitches itself as a safe place to make confessions and tell secrets. A more extreme anonymous social media, in which we can see the results of a lack of community guidelines in the design, is 4chan. By analyzing these platforms, we can better understand the effects that design decisions have on the way in which applications are used.

Background

When looking at the relationship between humans and anonymity, the concept of anonymity in online environments is best understood as a spectrum, divided into the categories of anonymity, pseudonymity, and real name (Nicole B. Ellison et al., 2016 p.1). Anonymity causes people to address a wider range of topics than in face-to-face settings, encompassing subjects that people feel would be awkward, embarrassing, or inappropriate to express through other channels (Nicole B. Ellison et al., 2016 p.7). This allows adolescents and teens to undergo important identity experimentation and social development (Teo Keipe et al., 2014 p.2). Unfortunately, this comes at the expense of enabling aggressive behaviors such as harassment and bullying within the anonymous online platforms (Kimberly Christopherson, 2006 p.4). Especially relevant in today's culture, true anonymity also enables the spread of misinformation.

In order to discourage the potential negatives associated with anonymity, software designers should utilize the Sociotechnical Systems Framework.

Socio-technical systems design methods are an approach to design that consider human, social and organizational factors, as well as technical factors in the design of organizational systems (Gordon Baxter et al., 2010 p.1). Gordon emphasizes the importance of a “human-centered design,” to best actualize a technology’s intended purpose. This calls for user participation in the development process to garner an explicit understanding of users, their tasks, and the environments in which they work (Gordon Baxter et al., 2010 p.4). Research into similar productions can provide a deeper understanding of problems, and the basis for potential solutions (Gordon Baxter et al., 2010 p.7). It is important to note that Baxter employs this line of thinking to try and maximize profit, however the same principles of analysis can be applied to maximize other goals of an application. Schaupp emphasizes the use of ethnographic analysis on technological structures to better understand resulting cultures (Simon Schaupp, 2021). Ethnography is the study of people and their interactions with cultural systems. Schaupp conducted research to better understand how customers interact with the differently designed retail websites (Simon Schaupp, 2021). He strongly emphasizes the impact of design on human interaction. By employing similar analysis tools to social media platforms, we find design decisions have a clear influence on the type of community culture the platform fosters.

Analysis of YikYak

There are many case studies of how past anonymous mobile applications have handled their disciplinary philosophies, and the effect this has had on the user base. One very popular example is the app YikYak. As Gordon Baxter suggested, to understand the politics of design we must first analyze the user base. YikYak was

designed to allow college students to connect and freely interact with their campus community, thus it was designed for young adults. (Vaterlaus, Mitchell J, 2017 p.2). YikYak experienced its initial growth in college towns (Erik W. Black et al., 2015 p.18). The task provided by YikYak is a fully anonymous message board that allows users to comment on, upvote, and downvote messages. This allows for informal research to be conducted by the user base as they see what gains the most likes, and what gets disliked. Caitlin Byrne concludes that anonymous social medias offer a new and unstudied source of data on critical issues like sex, class, and race (Caitlin Byrne, 2017 p.805). These topics have been extensively studied in public settings, but anonymous settings tend to allow people to respond more candidly (Caitlin Byrne, 2017 p.805). The data clearly demonstrates users' increased willingness to speak about personal topics when under the sense of anonymity. (John Bargh et al., 2002 p.46) Users posting questions anonymously are informally conducting personal research into what their local area thinks on a topic. That being said, the types of topics that get brought up and discussed vary largely from anonymous platform to platform, largely due to differences in behavior policies and enforcement. Analysis of YikYaks from large sample sizes offers us interesting information regarding anonymity and human behavior.

Despite its initial popularity, the app received backlash due to the large amount of cyberbullying occurring on the app. Erik W. Black and his colleagues speak to the rule changes that YikYak underwent in 2016 to increase accountability of accounts. First, YikYak began requiring cell phone number authentication to sign into YikYak and used these phone numbers as an account tracking mechanism (Erik W. Black et al. 2021). This gave them a means of tracking how many reported posts each account

accumulated over time, and to effectively blacklist users who had too many strikes (unless that person gets a new cell phone number, they would no longer have access to YikYak). YikYak employs a self-governing policy, allowing users to report posts, and removing posts and users who receive a reasonable number of reports. By doing a meta analysis of 4,001 YikYak posts, Black and his colleagues found that YikYak in general is not a venue for widespread anti-social behavior (Erik W. Black et al. 2021) . In terms of targeting specific people, they found that postings mentioning an individual's first or last name represented only 0.3% of the analyzed postings (Erik W. Black et al., 2015 p.20). While this could be for a variety of reasons, YikYak's banning policy was overall a functional and efficient method for community enforcement (Erik W. Black et al., 2015 p.21). By looking at the research that has already been completed in the world of anonymous social media, it is clear that the handling of disciplinary demerits in a quick and permanent manner would greatly discourage the improper use of the app, while still maintaining the benefits of anonymity. The change in YikYak's design from 2016 to the relaunch in 2021 lead to a much less toxic culture. This supports the concept that an effective enforcement policy of a thoughtful user guideline can have very tangible effects on the way users interact with one another.

Analysis of Whisper

Whisper is an anonymous social media app developed in 2012 and marketed as a safeplace to make confessions, share secrets, and post anonymously on different topics. After using the web application and making a few posts, I've learned that making a post first involves selection of a picture (you can use your own image or one of whisper's provided images). You then write text to overlay on the picture, select the font

type and the general topic of the post. Whisper has topics (similar in operation to reddit), and users are given randomly generated usernames to protect their identity. Unlike YikYak, Whisper gives users the ability to directly message the creator of a post, creating a personal message board between the two similar to Instagram's DM mechanism. The New York Times describes the target audience of Whisper as Gen Z (Alex Williams, 2015). Similar to YikYak, Whisper targets the teenage to young adult demographic. About 70% of the userbase is female, according to Whisper (J. V. Grove, 2014).

In addition to a crowdsourcing-based user reporting mechanism (similar to that of YikYak), Whisper also has dedicated employees to moderate whispers looking for obscenities (J. V. Grove, 2014). Wang and his colleagues compiled a complete 3-month trace of the Whisper network covering 24 million whispers written by more than 1 million unique users to better understand how anonymity and the lack of social links affect user behavior (Gang Wang et al. 2015). The first datapoint to note was that a surprising 18% of the whispers made during these three months were deleted, either by the user who posted it or by moderators (Gang Wang et al. 2015). Wang believes that the majority of deleted posts were done so by moderators, as the posts were often deleted past the common threshold of time for interaction. "Sexting" was the most common topic of these deleted whispers. About 25.4% of the users who posted at some point in the three months had a Whisper deleted, with about half of those users only having one whisper deleted (Gang Wang et al. 2015). A large portion of the whispers in the deleted dataset are repeats, leading Wang and his colleagues to believe that a user whose post got deleted would frequently regenerate a new random nickname and then repost the same

whisper (Gang Wang et al. 2015). This loophole helps to prevent the system from being able to flag their account with multiple strikes and banning them. This points to the greatest flaw in the Whisper enforcement policy; blacklisting of accounts is tracked by user nicknames, which can be easily regenerated.

A platform in which about 18% of the posts are deemed to be outside of the desired culture of the app is not effective. As mentioned prior to this, YikYak uses cell phone numbers to prevent reentry of a banned user into the YikYak system. By effectively removing banned users, YikYak has more control over the general behavioral culture of their platform. Also, the lack of a one on one direct message feature in YikYak helps limit the potential to abuse the app. Whisper's design decision to include a DM feature has led to some users soliciting hookups and showing pervertish tendencies (J. V. Grove, 2014). These are some examples of how design implementation, and choice of available features, directly impact the culture and usage of a platform.

Analysis of 4Chan

4chan is a fully anonymous message board that uses no usernames, has no DM feature, and operates using threads posted to topic boards, similar to reddit. Most interestingly, threads expire and delete themselves if no new messages are added for a certain period of time (Caitlin Dewey, 2014). The sense of impermanence within the threads further encourages anonymous users to say whatever is on their mind. The site's target audience isn't necessarily by design, but 4chan's own statistics suggest the "vast majority of its users are young, college-educated men with an interest in Japanese culture, video games, comics and technology" (Caitlin Dewey, 2014). It is important to

note that due to the level of anonymity offered by the platform, these statistics are by self reported (non-scientifically sound) surveys, although they still offer insight.

Policies vary for different discussion boards. Across all boards, users are banned for violating U.S. law, from posting other people's personal information, and impersonating site administrators (Caitlin Dewey, 2014). Beyond this, 4chan operates on two sets of rules, an extremely lenient rule set for the "random" board /b/, which allows nudity, racism, and other grotesque posts; and another rule set for the rest of 4chan's boards that do not allow for obscenities. The /b/ thread deletes threads very quickly, such that a website refresh every few minutes presents completely different results. This offers an extreme case study of complete anonymity within a platform with little to no rules or consequences. In terms of rule enforcement, 4chan has teams of referees it calls "moderators" and "janitors," who can delete posts and ban users (Caitlin Dewey, 2014). Users can also report posts, which flags it to moderators.

In order to understand the resulting app culture from an application with little to no rules, Bernstein and his colleagues completed an analysis of 5,576,096 posts in 482,559 threads of the /b/ board. They found that the median lifespan of a thread before it self-destructed was just 3.9 minutes, with the longest lasting 6.2 hours (Michael S. Bernstein et al. 2021). They found that not only does anonymity invoke disinhibition on /b/, but styling the collective as "Anonymous" also suggests de-individuation and mob behavior (Michael S. Bernstein et al. 2021). In other words, the design decision to list each user as "anonymous" (as opposed to giving them a randomly generated username like YikYak and Whisper do) increases the mob mentality and directly impacts the culture on the website. The content on the fully anonymous and ephemeral board /b/ is

vulgar to say the least. The design decision of the platform to have expiring threads “create a powerful selection mechanic by requiring content the community wants to see be repeatedly reposted, and potentially remixed” (Michael S. Bernstein et al. 2021). It encourages users to quickly reply to threads they’re interested, with less thought in their comments. Ephemerality of posts lead to a culture of less intelligent conversation, and an emphasis on hyperbole and quick memes. The extremity of 4chan’s /b board further shows evidence that platform design decisions can greatly impact the behavior and content produced by the users. It is clear that less regulations leads to a more toxic environment, and (for the most part) fails to produce fruitful communications and discussions of ideas.

Discussion

After analyzing these three anonymous social media platforms, it is clear there are noticeable differences in the cultures and way that users interact with one another. These behavioral differences are the result of different design decisions. By using Baxter’s and Schaupp’s analysis techniques we carried out an ethnographic analysis on different social media platforms to better understand the resulting user base’s cultures and interactions.

In 4chan’s case, we see the design decisions of user anonymity, ephemerality of threads, and light policy enforcement enabling a toxic environment within the platform. The lack of pseudo anonymous usernames contributes to further disrespect and misuse of the platform. Racism, extremism and pornography are extremely common in 4chan’s

/b/ board due to its design. In one of the few places on the internet where anything goes, users will take advantage of it and post heinously.

Whisper offers a unique perspective from a middle ground platform, where pseudo random usernames are assigned, and a user behavior policy is enforced. What Whisper lacks in its design is a permanence of discipline for violating the desired culture of the platform. This leads to repeat violators, as can be seen by the 18% of new posts being deleted. To fix this issue would sacrifice some level of anonymity, as historical user data would need to be stored to effectively ban specific users. YikYak does this through its login method, which only requires a cell phone number (which is verified at login by text). No further data is requested from the user (no name, age, etc.), however the cell phone number is used as a unique identifier of the account to give permanent punishments when needed. YikYak uses randomly generated usernames to avoid a mob mentality, and the developers chose not to include a DM feature to limit predatory capabilities. These all contribute to the revised YikYak culture we see since the relaunch.

Profanity can be used as one measurement of an ill-natured culture within a platform. When compared to a non-anonymous app, such as Twitter, YikYak has a similar level of profanity used. In an observational study, Martin Saveski and his colleagues found that 6.29% of all yaks (posts and replies), and 5.38% of all tweets, contain vulgar words (Martin Saveski et al. 2016). They also found that posts containing offensive language are more likely to be downvoted than not in YikYak, suggesting community-led censorship is a viable solution, even with anonymity (Martin Saveski et al. 2016). They point out that each user's "Yakarama" score helps to encourage positive

posting. Yakarma is a user's net total upvotes and downvotes that they've received on their posts (Martin Saveski et al. 2016). YikYak users enjoy boasting their Yakarma scores to their friends, and want to avoid making posts that would receive downvotes. This is yet another example of a design decision within the platform having a positive effect on upholding the desired culture.

Future Research

YikYak is local and anonymous, whereas Whisper and 4chan are global and anonymous. The lack of commonly used local anonymous social media to compare to YikYak is a shortcoming of this analysis, and would deserve further attention in future research. Apps such as leak, truth, secret3.0, and so on would be viable comparisons if they were to grow more popular, thus providing larger datasets to pull from. At this time, however, YikYak's popularity as a locally anonymous designed app is rather unique. On top of this, a meta-analysis of the differences in specific topic areas that tend to be brought up within each platform could offer further insight into the general culture of each platform. This could be used to confirm/disprove the advantage of offering anonymity in social media to facilitate discussion. From the research I've conducted, YikYak's policy revision was very effective in building a new behavioral culture within the platform, while still maintaining their initial vision, an anonymous communicative platform to connect local communities.

References

- Bargh, John A., McKenna, Katelyn Y., Fitzsimons, Grainne M. (2002). "Can you see the real me? Activation and expression of the "true self" on the Internet." *Journal of Social Issues*,
<https://www.researchgate.net/file.PostFileLoader.html?id=57c10f5ef7b67e73423ad5a1&assetKey=AS%3A399544220176387%401472270173879>
- Baxter, Gordon, Sommerville, Ian (2010). "Socio-Technical Systems: From Design Methods to Systems Engineering." Oxford Academic,
<https://academic.oup.com/iwc/article/23/1/4/693091>.
- Black, Eric W., Mezzina, Kelsey, Thompson, Lindsay A. (2015, December 18). "Anonymous social media – Understanding the content and context of Yik Yak." https://www.sciencedirect.com/science/article/pii/S0747563215302582?casa_token=TGvFdixyVoAAAAA%3A5h1yQBsO0e8Lyy6RxOGtPpqWEMPqgcZ3kCs6OKmf5eFgk-KYKpEV--Pe2y4GJDtZgRs45Hu2.
- Blackwell, Lindsay, Ellison, N. B., Lampe, Cliff (2016) "The Question Exists, but You Don't Exist with It: Strategic Anonymity in the Social Lives of Adolescents." *SAGE Journals*, <https://journals.sagepub.com/doi/full/10.1177/2056305116670673>.
- Byrne, Caitlin (2017) "Anonymous Social Media and Qualitative Inquiry: Methodological Considerations and Implications for Using Yik Yak as a Qualitative Data Source." *SAGE Journals*, Digital Spaces,
https://journals.sagepub.com/doi/pdf/10.1177/1077800417731081?casa_token=c-MVI7GZnasAAAAA%3AooAsBDCpERslegnmAl_WwxOha_Mf9IXeJYIYi104hpHMrijZagoPe0OVI5_Ye_wMC4B3fL0uoTs.
- Christopherson, Kimberly M. (2006) "The Positive and Negative Implications of Anonymity in Internet Social Interactions: 'on the Internet, Nobody Knows You're a Dog.'" *Science Direct*, Elsevier
<https://www.sciencedirect.com/science/article/pii/S0747563206001221>.
- Dewey, Caitlin (2014) "Absolutely everything you need to know to understand 4chan, the Internet's own bogeyman", *The Washington Post*,
<https://www.washingtonpost.com/news/the-intersect/wp/2014/09/25/absolutely-everything-you-need-to-know-to-understand-4chan-the-internets-own-bogeyman/>
- Gang Wang, Bolun Wang, Tianyi Wang, Ana Nika, Haitao Zheng, Ben Y. Zhao (2015) "Whispers in the Dark: Analysis of an Anonymous Social Network." *Proceedings of the 2014 Conference on Internet Measurement* Pages 137–150

<https://dl.acm.org/doi/abs/10.1145/2663716.2663728>,
<https://sites.cs.ucsb.edu/~tianyi/doc/whisper.pdf>

Grove, J. V. (2014) "*Secrets and lies: Whisper and the return of the anonymous app.*" CNet News,
<https://www.cnet.com/tech/services-and-software/secrets-and-lies-whisper-and-the-return-of-the-anonymous-app/>

Keipe, Teo, Atte Oksanen (2014) "*Self-Exploration, Anonymity and Risks in the Online Setting: Analysis of Narratives by 14–18-Year Olds.*" Taylor & Francis Online, Journal of Youth Studies,
<https://www.tandfonline.com/doi/full/10.1080/13676261.2014.881988>.

Michael S. Bernstein, Andres Monroy-Hernandez, Drew Harry, Paul Andre, Katrina Panovich and Greg Vargas (2021) "*4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community.*"
<https://www.media.mit.edu/publications/4chan-and-b-an-analysis-of-anonymity-and-ephemerality-in-a-large-online-community/>

Saveski, Martin, Chou, Sophie, Roy, Deb (2016) "*Tracking the Yak: An Empirical Study of Yik Yak.*" International AAAI Conference on Web and Social Media, North America.
<https://dspace.mit.edu/bitstream/handle/1721.1/125823/yikyak.pdf?sequence=2&isAllowed=y>

Schaupp, Simon, *Technopolitics from Below: A Framework for the Analysis of Digital Politics of Production*. Nanoethics 15, 71–86 (2021).
<https://doi.org/10.1007/s11569-021-00386-8>

Vaterlaus, Mitchell J. (2018) "*Yik Yak: An Exploratory Study of College Student Uses and Gratifications.*" SAGE Journals, UNF,
<https://journals.sagepub.com/doi/pdf/10.1177/0270467617738902>.

Williams, Alex (2015) "*Move Over, Millennials, Here Comes Generation Z*". The New York Times,
<https://www.nytimes.com/2015/09/20/fashion/move-over-millennials-here-comes-generation-z.html>