

4chan and /b/:

An Analysis of Anonymity and Ephemerality in a Large Online Community

Michael S. Bernstein¹, Andrés Monroy-Hernández¹, Drew Harry¹,
Paul André², Katrina Panovich¹ and Greg Vargas¹

¹Massachusetts Institute of Technology
Cambridge, MA 02139

{msbernst, amonroy, dharry, kp, gvargas}@mit.edu

²University of Southampton
SO17 1BJ, United Kingdom
pa2@ecs.soton.ac.uk

Abstract

We present two studies of online ephemerality and anonymity based on the popular discussion board /b/ at 4chan.org: a website with over 7 million users that plays an influential role in Internet culture. Although researchers and practitioners often assume that user identity and data permanence are central tools in the design of online communities, we explore how /b/ succeeds despite being almost entirely anonymous and extremely ephemeral. We begin by describing /b/ and performing a content analysis that suggests the community is dominated by playful exchanges of images and links. Our first study uses a large dataset of more than five million posts to quantify ephemerality in /b/. We find that most threads spend just five seconds on the first page and less than five minutes on the site before expiring. Our second study is an analysis of identity signals on 4chan, finding that over 90% of posts are made by fully anonymous users, with other identity signals adopted and discarded at will. We describe alternative mechanisms that /b/ participants use to establish status and frame their interactions.

Introduction

Identity representation and archiving strategies are central features to the design of online communities. However, our current understanding of them focuses mainly on strong identity and permanent archival. Researchers and practitioners argue that real names and pseudonyms can help “promote trust, cooperation, and accountability” (Millen and Patterson 2003), whereas anonymity may make communication impersonal and undermine credibility (Hiltz, Johnson, and Turoff 1986; Rains 2007). Influential industry players like Facebook argue that pseudonyms and multiple identities show “a lack of integrity” (Kirkpatrick 2010). Similarly, data permanence is also the norm: search engines will resurface content years after it is created (Rosen 2010), social network sites allow friends to browse updates and photos from years ago, and online communities will often expect newcomers to read their archives (Millen 2000). Some scholars have questioned these design approaches, suggesting that anonymous contributions and ephemeral participation online can be desirable (Lampe and Resnick 2004;

Ren, Kraut, and Kiesler forthcoming; Grudin 2002). However, we have a limited understanding of how an anonymous and ephemeral community design might actually play out — especially at large scale.

In this paper we analyze one such large-scale, anonymous, and ephemeral community: the imageboard website 4chan. We focus on 4chan’s first and most popular board, the “random” board known as /b/. Our goal is to use /b/ as a lens to understand the concepts of *anonymity* and *ephemerality* online. /b/ implements these concepts in more extreme ways than most other online communities. First, posts are fully anonymous by default and very rarely contain pseudonyms or other identity signals. This lack of identity makes traditional reputation systems unworkable. Second, instead of archiving conversations, /b/ deletes them when newer content arrives — often within minutes — which leads to a chaotic, fast-paced experience. By making complete anonymity and content deletion the norm, /b/ lets us study these concepts *in situ* at a larger scale than before.

This work quantifies the outcomes of 4chan’s design decisions, and starts a discussion on how those decisions affect the community and its culture. Although /b/’s implementation of anonymity and ephemerality are extreme and unusual, the concepts themselves are not unique. Sites like Twitter, for instance, feel ephemeral because of their continuous content stream, while others like Formspring use anonymity as a core feature. By studying the impact of different points on the identity and archival continuums, we can broaden our understanding of community design strategies.

Readers may know of 4chan and /b/ (Figure 1) from their influence on Internet culture and media coverage of their off-site activities. The site boasts over seven million users (Poole 2010) and is a prolific meme factory: it originated popular memes like LOLcats (Rutkoff 2007) and rickrolling (Leckart 2009). Additionally, some 4chan and /b/ members have been known to participate in highly visible off-site activities. These activities include manipulating a Time Magazine poll to elect 4chan’s creator the “World’s Most Influential Person” and participating in hacktivist group ‘Anonymous’. Anonymous has executed highly visible protests of the Church of Scientology (Coleman 2011) and DDoS attacks against Mastercard and Paypal in support of Wikileaks (Mackey 2010). Reactions are diverse: while memes have brought the site positive media attention (Brophy-

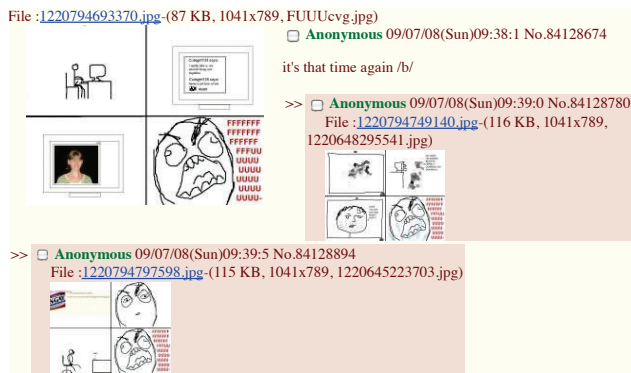


Figure 1: A “rageguy (FFFFUUUU) comic” themed thread on /b/. Source: 4chanarchive.org, where /b/ members save memorable threads.

Warren 2008), harassment and activism have often been negatively and sensationally covered. For example, a Fox News affiliate called 4chan the “internet hate machine” (Shuman 2007). However, some media outlets have profiled the site in more nuanced ways (Dibbell 2009; 2010; Poole 2010). Much of this coverage is a response to /b/’s distinctive culture – a culture that merits critical analysis beyond the scope of this paper. Here, we focus on the board’s design choices of anonymity and ephemerality, and how they may support or influence its culture.

In this paper, we perform a content analysis and two data-driven studies of /b/, focusing on anonymity and ephemerality. To begin, we survey related work. We then introduce 4chan, its design, and the /b/ board. To ground our discussion of the site, we perform a content analysis on a sample of /b/ threads. We then turn to our two studies: 1) ephemerality, tracking the site’s tempo and content deletion dynamics; 2) anonymity, examining participant practices around identity. A note before proceeding: large portions of the 4chan site, and /b/ in particular, are offensive or obscene. We warn that quotes and vocabulary in this paper may offend.

Related Work

Our work builds on prior literature on anonymity and ephemerality. 4chan and /b/ can contribute insights into how this literature plays out in the wild at large scale.

Online communities choose points on the spectrum of anonymity — from completely unattributed to real names. For example, while Facebook embraces real names (Facebook 2010), Myspace does not (Dwyer, Hiltz, and Passerini 2007), and some Usenet boards allow posting from anonymous e-mail addresses (Donath 1999). Slashdot decided to enable anonymous commenting so users could feel more free to speak their minds, then controlled behavior with user moderation of comments (Lampe and Resnick 2004). However, while they may allow fully anonymous posting, anonymity is much less common in these communities than on /b/. Gómez et al (2008) found that fully anonymous posts made up only 18.6% of Slashdot comments. Instead, pseudonymity tends to become the norm as usernames allow members to build a reputation. Our work extends this research by studying the dynamics of a more thoroughly anonymous community.

Evidence is mixed on how anonymity may affect an on-line community. In many scenarios, researchers argue for the importance of identity-based reputation systems in promoting pro-social behavior (Millen and Patterson 2003). However, anonymity may foster stronger communal identity, as opposed to bond-based attachment with individuals (Ren, Kraut, and Kiesler forthcoming). Anonymity may impact participation: it increases equity in classrooms (Collins and Berge 1995), but results in more ‘flaming’ on e-mail lists (Thompson and Ahn 1992). 4chan and /b/ play out these concepts on a larger stage than has been previously studied.

Computer-mediated communication has studied anonymity in small groups, and our work reconsiders their results in larger online communities. Removing traditional social cues can make communication impersonal (Short, Williams, and Christie 1976) and cold (Hiltz, Johnson, and Turoff 1986), and choosing to remain anonymous will undermine credibility (Rains 2007). However, we will argue that /b/’s community has developed alternative credibility mechanisms — via language and images — that still function effectively. Anonymity can also have positive outcomes: groups working anonymously and with critical confederates produce more ideas (Jessup, Connolly, and Galegher 1990); non-anonymous groups feel more personal, but have less overall cohesion (Tanis and Postmes 2007).

Ephemerality is rare in a large-scale online community, and to our knowledge, we are the first to study it directly *in situ*. Most communities that have been studied rely heavily on archives. For example, Millen (2000) reports that community members often expect each other to search group archives before asking new questions. Ephemerality may have community-wide downsides – a lack of history tends to decrease cooperation in social dilemma games (Fehr and Gächter 2000). However, instituting permanence in previously-unarchived chat rooms has elicited strong negative reactions (Hudson and Bruckman 2004).

Through our investigation of /b/, we hope to contribute to scholarly conversations about data permanence. For example, Grudin (2002) suggests that we evolved to live in an ephemeral world, yet our technology takes us from the “here and now” to the “everywhere and forever.” Similarly, Mayer-Schönberger (2009) emphasizes the value of “societal forgetting,” where “the limits of human memory ensure that people’s sins are eventually forgotten.” Blanchette and Johnson (2002) pointed to the recognition of social forgetfulness in three areas of social policy (bankruptcy law, juvenile crime records and credit reports), arguing that “data retention and disposal should be addressed as a fundamental characteristic of information.”

These topics are not just of academic interest, but have clear practical implications for online social environments. Practitioners face similar challenges. For example, online game retailer Blizzard recently reversed a decision to require the use of out-of-game identities in its online forums (M.G. 2010), Formspring found both popularity and controversy by allowing teens to ask anonymous questions of their friends (boyd 2010a), and AOL angered its users when a hacker de-anonymized old search logs (Barbaro and Jr. 2006).

Type	%	Description	Example
Themed	28%	Setting theme, often with an exemplar image.	<i>ITT ["In this thread"] we only post stuff we have laughed at so hard we had tears</i>
Sharing content	19%	Offering content for the community to enjoy or critique.	<i>This guy is a hero. :) http://www.youtube.com/[xxxxx]</i>
Question, advice and recommendation	10%	Asking for suggestions or, often quite intimate, life advice.	<i>Soup /b/ Recently I've been hanging out with a girl a lot, we're both in college. I spend the night at her place all the time and we kiss and whatnot. Problem is, she just broke up with her ex [...] and I know she's not over him. I really like her but I dunno what to do, so what do /b/?</i>
Sharing Personal Information	9%	Sharing or requesting content with personal information.	<i>U JELLY? ["You jealous?"] This is me suiting up at my formal looking fucking brilliant, then there is you fags sitting back and watching.</i>
Discussion	8%	Calling for discussion, debate or some back-and-forth over a topic.	<i>Hi Anonymous! So ive started this game called League of Legends a few days ago. [...] Is there anyone here who also plays it? Lets talk about it!</i>
Request for item	8%	Requesting information about previously-seen images, or other valuable information such as credentials for pay sites.	<i>anyone has a pic of that star wars battle tank with the german insignia shopped ["photoshopped"] on it? in return tits [a misogynous but common mechanism for "paying back" a favor with pornographic images]</i>
Request for action	7%	Intending to agitate for real-life action, like harassing another website	<i>Make a group saying [name] is awesome on face book. DO IT FAGGOTS</i>
Meta	5%	Discussing /b/ itself or playing with the site's mechanics (e.g., post numbers)	<i>Heidi Ho there /b/ I'm a Newfag and now that i've been here all Summer I was wondering if i need a letter of recommendation From a Registered OldFag?</i>
Other	6%	Unable to categorize.	<i>excuse all the blood</i>

Table 1: Content typology of threads on /b/: appearance frequency and an exemplar (some quotes are paraphrased). ($n = 598$)

4chan and /b/

4chan was created in 2004 by Christopher Poole as an online discussion board focused on Japanese anime (Sorgatz 2009). It has grown from its anime roots to encompass sixty boards on topics ranging from politics to fashion, science and “sexy beautiful women.” Poole, better known by his pseudonym *moot*, created 4chan by copying the format of Futaba Channel, a popular Japanese discussion forum.¹ 4chan’s aesthetic is simple, though it can appear confusing and cluttered: the Wall Street Journal describes it as “archaic [...] a quaint throwback to the earliest webpages” (Brophy-Warren 2008).

4chan is composed of boards, threads and posts. Each board is themed (e.g., /v/ is “Video Games”). Like most discussion boards, 4chan groups posts into threads (Figure 2). Posts starting a thread are required to include an image while images on replies are optional. Threads are organized into pages, where each page previews fifteen threads with their original post and a small sample of replies — users can click through to read the entire thread.

In this paper, we focus on the 4chan “random” board, known as /b/.² We focus on /b/ not only because it is 4chan’s first and most active board — it claims 30% of all 4chan traffic — but also because, in the words of its creator, it is the “life force of the website”, and the place where “rowdiness and lawlessness” happen (Sorgatz 2009).

Content ephemerality on 4chan is enforced by thread expiration and a large volume of incoming content. Threads begin on page one and are pushed down as new threads are

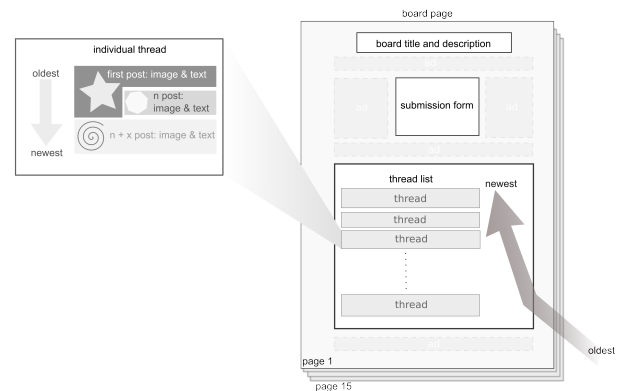


Figure 2: Structure of a 4chan board.

added. If a user replies to a thread, it is *bumped* back to the top of the first page. If the thread reaches the bottom of the fifteenth page, the thread is removed permanently and its URL returns a ‘Page Not Found’ error. This entire process can take place over a matter of minutes, as Study 1 will demonstrate.

4chan’s anonymity plays out through its posting mechanisms and defaults. Unlike other sites, where being anonymous usually means not using your “real” name or identity, most posts on /b/ are disconnected from *any* identity. There are no accounts; all information is entered on a per-post basis. If the user does not change the default empty name field, 4chan assigns the name *Anonymous*. Even if a user claims a pseudonym, any other user can claim it themselves in any subsequent post. 4chan does have a cryptographically-

¹<http://www.4chan.org/faq#what4chan>

²<http://boards.4chan.org/b/>

powered feature for users to guarantee their identity, called a *tripcode*. Tripcodes use a password to generate a unique string after the username (e.g., *username!!Oo43raDvH6l*), giving the password-holder a unique and inimitable identifier. However, 4chan users largely eschew tripcodes and pseudonyms, as Study 2 will show.

Content posted on /b/

/b/'s content is frequently intentionally offensive, with little held sacred. There is racist, sexist, homophobic language, groups are often referred to using a "fag" suffix (e.g., new members are "newfags", British users are "britfags"), and a common response to any self-shot picture by a woman is "tits or GTFO" (post a topless photo or get the f*** out). This language is part of the group identity: pushing the bounds of propriety in order to "hack the attention economy" and turn heads (boyd 2010b). While the content on /b/ can be offensive, it can also be funny, open, and creative, as its creation and promotion of numerous memes attests to.

In order to characterize the content and discourse on the board, we began with eight months of participant observation on the site. Using a series of informal samples of thread-starting posts on /b/, we conducted a grounded analysis (Charmaz 2006) similar to that which has been applied to other kinds of online participation (Naaman, Boase, and Lai 2010). After a few rounds of iteration, we settled on a scheme with nine high-level categories to describe the different kinds of posts that initiate threads on /b/. To measure the relative frequency of these thread categories, we collected a sample of 598 thread-initiating posts over the course of ten days (November 16–26, 2010). The threads were selected such that their temporal distribution matches the underlying distribution of 4chan posts. Our sample included the text and accompanying image from the post that started the thread.

Table 1 reports the thread composition in our sample. In keeping with /b/'s identity as an image board and the requirement that each thread-starter post an image, a common purpose of the board is to share images and web content. The two most prevalent thread types (Themed, 28%, and Sharing Content, 19%) both revolve around images and make up nearly half of all threads in our sample. There is also evidence for the off-/b/ activities that the media focuses on: threads attempting to organize such activities (Request for Action) make up 7% of the sample. Of those threads, the poster often attempted to generate comments on Facebook or Youtube pages or get people to call a particular phone number. /b/'s posters often disdain such calls because they are seen as self-serving, dismissing them by replying "/b/ is not your personal army." In the future, we plan to examine whether content category is associated with outcomes like thread length or reply frequency.

Study 1. Ephemerality

Ephemerality is one of /b/'s most striking qualities. The board moves at such a fast pace, and threads expire so quickly, that the site is largely different with each page refresh. In this section, we quantitatively describe the temporal properties and dynamics of posts on /b/. We compare /b/'s

volume to other sites to provide a sense of turnover speed. We then relate how /b/ users have developed coping mechanisms like personal archives to cope with the quick expiration of material, as well as practices like "bumping" to keep threads alive. Finally, we explore how strict ephemerality policies may actually encourage increased community participation.

Method

We collected a dataset of activity on /b/ for two weeks: July 19–August 2, 2010.³ This data includes 5,576,096 posts in 482,559 threads. Although there are likely some phenomena that influence /b/'s posts over longer time scales than two weeks (e.g. holidays), most of the major daily and weekly cycles are represented in a sample of this size. The dataset is missing a negligible number of posts during high-load periods. These missing posts are relatively randomly distributed and we do not believe that they impact our analysis. We did not capture images due to concerns over the nature of the material that may be posted to the website.

We calculated the time each thread spent on /b/ by replaying the history of all post events from our two-week dataset. We used creation timestamps as reported by the website to simulate the positions of each thread. For example, after a thread has a new reply, it moves to the top position on the first page; after a post is made in another thread, the first thread is pushed down to the second position. By replaying this history, we calculated the lifetime for each thread.

Results

Entire Lifetime The majority of threads have a short lifespan and a small number of replies; the median life of a thread is just 3.9 minutes. Thread lifetimes are right-skewed similar to a power law, making the mean less meaningful: 9.1 minutes ($\sigma = 16.0$ min). The fastest thread to expire was gone in 28 seconds (i.e., a thread with no responses during a very high activity period); the longest-lived lasted 6.2 hours (i.e., a thread with frequent new posts to bump it). Six hours is a very long time in /b/, but it is near-instantaneous when compared to the forever-archived nature of most other websites.

The longest-lived thread in our dataset was a discussion of paganism. The original poster was a pagan who advertised the opportunity to Ask A Pagan Anything, remarking "go on do your worst (or alternatively actually get my respect and actually ask something useful)." One question was "how do you worship your so called gods?", with the answer "From day to day just by reveling in the beauty and wonder of life, in all it's forms [sic] from studying martial arts to sat-iffie [sic] my masculinity to taking care of my garden." Other questions included "How does it feel knowing Christianity raped your religion?", "What exactly do you worship?", and "Do you believe in magic?" Other long-lived threads fell in the "themed" category, like a "creepypasta thread", "info threads" (posting useful knowledge about some topic, like how to tie a tie or keyboard shortcuts), and self-shot nudity.

³More information about our data collection tool available at <http://projects.csail.mit.edu/chanthropology>

Short-lived threads on /b/ varied, but many were failed attempts to get the community's attention (e.g., "Well guyz, I hope you're glad the captcha is gone, woo yeah, let's all be random lolz!"). The short-lived posts often came in spurts during high-activity periods when it was easy to miss them.

As might be implied by short lifetimes, a large number of threads (43%) get no replies at all; the median is 2 posts per thread, the original post and one reply. This 43% figure is roughly consistent with Usenet, where 40% of posts get no reply (Joyce and Kraut 2006). Again, some threads become quite large, resulting in a mean of 13.27 posts/thread ($\sigma = 37.28$ posts, min = 1, max = 519).

First Page Only Another way to look at how ephemerality plays out in the board is by looking at each thread's exposure to the first page. The first page of results is where many items, like search results, get much of their overall visibility and click-throughs (Joachims et al. 2005).

The median thread spends just 5 seconds on the first page over its entire lifetime. The mean time on the first page is 36 seconds ($\sigma = 109$ sec). The fastest thread was pushed off the first page in less than one second (actually, 58 of them shared this dubious honor), and the most prominent thread spent 37 minutes on the first page cumulatively over its lifetime.

The thread that spent the most time on the first page was a "roll" thread (a meta thread in our content analysis, playing with the mechanics of the board). In a roll thread, /b/ posters reply to get a 4chan-assigned post number (e.g., 1234567), and the last digit of their post number instructs them on an action to take. In this case, participants had to share personal secrets, a game that combines Spin the Bottle with Truth or Dare. Example responses included: "honestly, i dont know. Nothing has ever made me feel too terrified for my life. Maybe when I had taken too much cocaine and thought my heart was gonna overwhelm itself", "I gave into temptation by having two boyfriends at once instead of 1", "chose 9 i get angry at the idea of people having control over me and abusing it or manipulating me, a weird contrast to my mind control fetish", and "blue 5: my biggest regret is not asking her out".

User Control over Ephemerality: Bumping and Sage 4chan has developed two main ways for users to control thread ephemerality: *bumping* and *sage*. *Bumping* means replying to a thread to keep it alive, sometimes explicitly with a phrase like "bump", "bumping" or "bamp". In our sample, we observed that 2.16% of all replies contained these words or similar inflections. This is a lower-bound estimate, since any post will effectively "bump" a thread. The second method of control is *sage*, which allows a user to comment on a thread but not bump it to the first page (i.e., bury it). This lets users comment on a disliked thread without attracting attention to it, and to count the reply toward a system-enforced bump limit, thus ensuring the thread will expire more quickly. We found that 0.77% of all replies used the sage feature. So, /b/ posters will explicitly manipulate the ephemerality of some threads, though they seem more likely to promote threads than bury them.

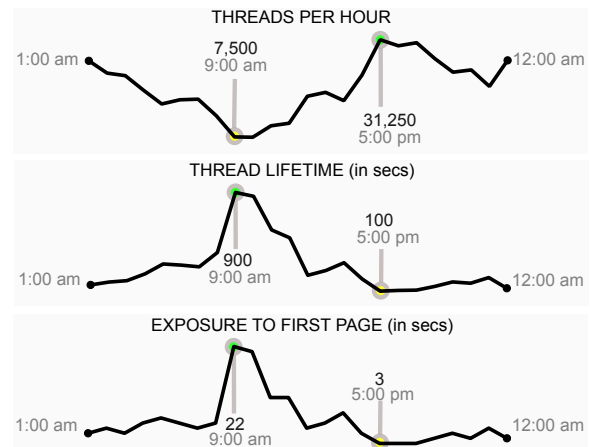


Figure 3: Daily board activity in our two-week dataset. Thread Lifetime and Exposure to First Page are medians; Threads per Hour averages the volume during that hour over fourteen days. All times in EST.

Time of Day /b/'s daily trends can help us understand how ephemerality is affected by time of day. The slower the traffic on /b/, the longer a thread will last.

Threads last the longest between 9am and 10am EST and expire fastest between 5pm and 7pm EST. High activity is sustained until 3am or 4am EST. This result suggests that, despite the not infrequent references to European and British users (e.g., "eurofags" and "britfags"), the demographics of /b/ are primarily North Americans that use the website after business or school hours. Figure 3 shows the synchronized spikes in lifespan and drop in number of posts per thread.

Discussion

/b/ creates a sense of ephemerality through a fast tempo and content deletion. In regards to tempo, we found that the board had roughly 35,000 threads and 400,000 posts per day. For comparison, Usenet volume (which still continues to grow) across all Big-8 newsgroups is 25,000 posts per day⁴, or 1/16th of /b/'s volume. Szabo and Huberman (2010) found that Digg has about 7,100 "threads" each day (1.3 million over six months), and 65,000 new YouTube videos each day. So, /b/ has roughly the same amount of posting activity as arenas like Usenet and Youtube, but *all of this activity happens in one forum board*.

Though the site may be ephemeral, /b/ users have developed other mechanisms to keep valuable content. For example, users often refer to having a "/b/ folder" on their computers where they preserve images for future enjoyment or remixing. /b/ posters ask others to dig into their archives; for example, this user wants to shock a friend, and donates an image of a cat in return: "Have a friend here, need the most fucked up shit you have in your /b folder, can't provide much, considering not on my comp, but here is a cat." /b/ users have also developed sites like 4chanarchive.org to save particularly important or "epic" threads.

Content deletion may play a role in pushing the /b/ community to quickly iterate and generate popular memes like LOLcats, Advice Dog and Archaic Rap. Having no history

⁴<http://www.newsadmin.com/top100tmsg.asp>

moderates some the “rich get richer” phenomena (Barabási and Albert 1999). 4chan’s founder has argued that /b/’s “lack of retention [...] lends itself to having fresh content,” so only the fittest memes survive (Sorgatz 2009). To keep content around, users must make an explicit decision to save it to their hard drive and repost it later.

Finally, and perhaps unintuitively, ephemerality may raise community participation. One may think users would see no point to contributing if their actions will be removed within minutes. However, if /b/ users want to keep a thread from expiring within minutes, they need to keep conversation active. This “bump” practice, combined with a norm of quick replies, may encourage community members to contribute content. This hypothesis was derived from our observations, and will need to be tested more rigorously.

Study 2. Identity and Anonymity

/b/ makes it easy to participate without requiring a real or even pseudonymous identity. In this section, we investigate the frequency of the most common names used by /b/ posters, we discuss the impact of anonymity on /b/’s culture, and report on /b/’s alternative status and authenticity signals. We also discuss the effects of anonymity-fuelled disinhibition, like those seen in the “relationship advice” threads and “Anonymous” raids.

Method

Using our two-week data sample, we analyzed the identity metadata of each post. /b/ allows posters to enter no name (“Anonymous”), choose any name, or use a cryptographic identity mechanism known as a tripcode. We investigated the prevalence of each of these identity markers in our dataset.

Results

It is extremely uncommon to post using a name or pseudonym on /b/. In our sample, 90.07% (5,022,149) of posts were credited to the default name “Anonymous” (Table 2). The closest comparison available in the literature is that anonymous commenting makes up 18.6% of Slashdot comments (Gómez, Kaltenbrunner, and López 2008). The remaining 10% use a wide diversity of names. Some relate to an inside joke where many users claimed one name, David,⁵ and others show mistaken uses of 4chan-specific keywords like *sage* or *noko*. Some users claimed to be “OP” (the original poster of the thread), demonstrating a way in which /b/ posters fluidly claim identity when needed.

E-mails are even less common (Table 3). Fully 98.3% (5,478,573) of posts in our sample did not contain an e-mail. Of those that did complete the e-mail field, 40.73% (39,725) are not actual emails but rather posts trying not to bump the thread (using the *sage* feature). The rest are misspellings of 4chan’s special commands, some of them temporary. For example, “:stopsound:” got rid of the vuvuzela buzzing that 4chan administrators added to the page during the 2010 World Cup.

⁵http://encyclopediadramatica.com/Operation_/b/ipolar

Name	Num. of Posts	% of Posts
Anonymous	5,022,149	90.07%
David	59,320	1.06%
–	14,070	0.25%
OP	13,576	0.24%
(blank)	13,077	0.23%
sage	13,003	0.23%
anonymous	6,150	0.11%
noko	5,727	0.10%

Table 2: The most popular post names on /b/ during our two-week sample. Anonymous has over ninety percent of the posts ($n = 5,576,095$).

E-mail	Num. of Posts	% of Posts
(blank)	5,478,573	98.25%
sage	39,725	0.71%
Noko	5,037	0.09%
:stopsound:	3,377	0.06%
:soundoff:	1,627	0.02%

Table 3: The most common emails on /b/ during our two-week sample. Most posts leave the e-mail field blank ($n = 5,576,095$).

Tripcodes are the only way a 4chan user can guarantee that they are the same author of a previous post; however, they are very rarely used. Five percent of posts (281,367) – one twentieth of our sample – contained a tripcode. This number is likely inflated, because participating in the “David” in-joke mentioned above required using a shared tripcode among many users, which is uncommon. Ignoring the “David!4changtcqk” tripcode occurrences lowers the total to 4% (211,068).

Discussion

The usual narrative around anonymity suggests that communities benefit by revealing participants’ names and reputations (Millen and Patterson 2003), and that anonymity will be a negative influence due to the “online disinhibition effect” (Suler 2005). Certainly, /b/ is a crude place and is given to antisocial behavior. Not only does anonymity invoke disinhibition on /b/, but styling the collective as “Anonymous” also suggests de-individuation and mob behavior. It may be safe for /b/ posters to act in a way they never would do offline because they can be relatively certain that their actions will not come back to haunt them.

However, the dynamics on 4chan and /b/ also suggest ways that anonymity can be a positive feature for communities. Disinhibition can be beneficial: in advice and discussion threads, anonymity may provide a cover for more intimate and open conversations. For example, in Table 1, the poster asks anonymously for advice about a potential girlfriend. Such threads are quite common. In addition, anonymity may encourage experimentation with new ideas or memes. As seen in Study 1, failure is quite common on 4chan: almost half of all threads receive no replies. Anonymity masks that failure, softening the blow of being ignored or explicitly chastised for trying to start uninteresting threads (Dibbell 2010). In communities with stronger identity mechanics, a history of poor posts would never go away. On 4chan, it is irrelevant.

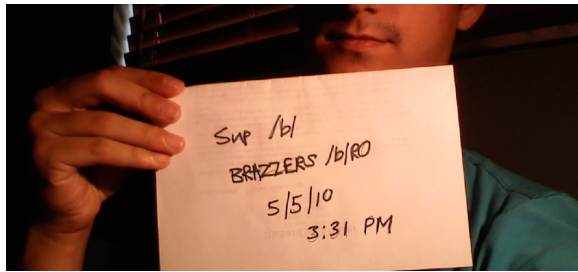


Figure 4: Timestamping lets users circumvent technological barriers to guarantee identity. Source: 4chanarchive.org.

/b/ has given rise to more fluid practices to signal identity and status in spite of, or perhaps because of, the lack of technological support. Because anyone can post a picture and claim to be that person, /b/'s posters have developed a practice of "timestamping" to guarantee authenticity. To claim identity, users often take a picture of themselves with a note containing the current day and time (Figure 4).

To communicate high status in the community, most users tend to turn to textual, linguistic, and visual cues. In many communities, including /b/, slang plays a role in delineating group membership (Eble 1996). Simply writing in 4chan dialect is non-obvious to outsiders and in-dialect writing serves as an entry-level signal of membership and status. Second, images on 4chan function a bit like fashion. Specific classes of images have periods of limited experimentation, turning into wider adoption, followed by subsequent abandonment. Fluency in the styles that are in vogue is an important way to signal status, as in fashion (Simmel 1957). Lack of fluency is dismissed with the phrase "LURK MOAR", asking the poster to spend more time learning about the culture of the board.

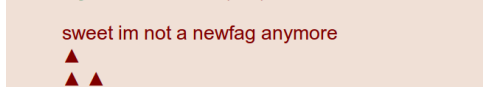
One example status signal in /b/ is the classic barrier for newcomers called "triforcing." Triforcing means leaving a post using Unicode to mimic the three-triangle icon of popular video game The Legend of Zelda:

□ Anonymous 08/30/09(Sun)17:59:38 No.156666179



Newcomers will be taunted by a challenge that "newfags can't triforce." Uninitiated users will then copy and paste an existing triforce into their reply. It will look like a correct triforce in the reply field; however, after posting, the alignment is wrong:

□ Anonymous 07/03/10(Sat)10:32:29 No.247423XXX



The only way to display high status and produce a correct triforce on 4chan is to use a complicated series of Unicode character codes. In signaling theory terms, we can think of triforcing as an index: a signal whose presentation is only possible by someone with particular skill or knowledge (Smith and Harper 1995).

That communities enforce boundaries and communicate status using language and differentiated social practices is perhaps not surprising in the social psychology literature. What we see as particularly noteworthy in this case is how

these boundary-sustaining practices are informed by the technical context in which they take place. Furthermore, the extreme nature of community practices on /b/ can obscure the underlying role these behaviors play to the casual observer. We see our role as partly one of translating these practices into terms familiar to scholars.

Conclusion

In this article we investigate the 4chan /b/ board as a vehicle for understanding the effects of ephemerality and anonymity in online communities.

Analyzing ephemerality via two weeks of site activity, we found that the median thread spends just five seconds on /b/'s first page before being pushed off by newer posts, and that the median thread expires completely within five minutes. Even in a world informed by Twitter and newsfeeds, where content is out of users' attentional sphere quickly, we argue that such rapid content *deletion* drives many of /b/'s community dynamics. On /b/, ephemerality and deletion create a powerful selection mechanic by requiring content the community wants to see be repeatedly reposted, and potentially remixed. We believe this is critical to the site's influence on internet culture and memes.

We then examined anonymity on /b/. We found that over 90% of posts are made completely anonymously, and that 5% of posts use system mechanisms like tripcodes to guarantee identity. Instead, the /b/ community uses non-technical mechanisms like slang and timestamping to signal status and identity. Consistent with common identity theory (Ren, Kraut, and Kiesler forthcoming), /b/'s anonymity is likely shaping a strong communal identity among a very large set of individuals.

We hope to open the door to future 4chan work. 4chan is widely credited with being the source of many online memes, making it an excellent venue for studying innovation diffusion. By tracking images (and genres of images) we could see how trends spread through 4chan and into the wild. It is also clear that 4chan represents only one part of a larger online ecosystem. Future work might focus on how 4chan's users move between other tools and interaction venues to organize both online and offline action. Finally, a closer study of the content on 4chan and its users would enable us to make more substantial claims about the relationship between 4chan's design and its users' practices and culture.

As large Internet players like Facebook or Google evolve their models for identity and archiving, it becomes increasingly important to understand what happens in large communities that occupy the opposite positions on the user identity and data permanence design continuums. Communities like 4chan have immense impact on Internet culture, and /b/'s anonymous, ephemeral community design is playing a strong role in that cultural influence.

Acknowledgements

We would like to thank for their feedback: danah boyd, Amy Bruckman, Gabriella Coleman, Danyel Fisher, Benjamin Mako Hill, Alex Leavitt, Rob Miller, Lisa Nakamura, Chris Schmandt, Christina Xu and Sarita Yardi.

References

- Barabási, A., and Albert, R. 1999. Emergence of scaling in random networks. *Science* 286(5439):509.
- Barbaro, M., and Jr., T. Z. 2006. A Face Is Exposed for AOL Searcher No. 4417749. *New York Times*.
- Blanchette, J., and Johnson, D. G. 2002. Data retention and the panoptic society: The social benefits of forgetfulness. *The Information Society: An International Journal* 18(1):33.
- boyd, d. 2010a. Harassment by q & a: Initial thoughts on form-spring.me. <http://dmlcentral.net/blog/danah-boyd/harassment-qa-initial-thoughts-formspringme>.
- boyd, d. 2010b. "for the lolz": 4chan is hacking the attention economy. <http://www.zephoria.org/thoughts/archives/2010/06/12/for-the-lolz-4chan-is-hacking-the-attention-economy.html>.
- Brophy-Warren, J. 2008. Modest web site is behind a bevy of memes. *Wall Street Journal*.
- Charmaz, K. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London: Sage Publications.
- Coleman, G. 2011. "'I did it for the lulz! but i stayed for the outrage:' anonymous, the politics of spectacle, and geek protests against the church of scientology." <http://techtv.mit.edu/videos/10237>.
- Collins, M., and Berge, Z. 1995. Introduction: Computer-mediated communications and the online classroom in higher education. In Berge, Z., and Collins, M., eds., *Computer mediated communication and the online classroom: Vol. 2. Higher education*. Hampton Press. 1–10.
- Dibbell, J. 2009. The Assclown Offensive: How to Enrage the Church of Scientology. *Wired*.
- Dibbell, J. 2010. Radical Opacity. *Technology Review*.
- Donath, J. 1999. Identity and Deception in the Virtual Community. In Kollock, P., and Smith, M., eds., *Communities in Cyberspace*. London: Routledge.
- Dwyer, C.; Hiltz, S.; and Passerini, K. 2007. Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. In *Proc. of AMCIS*.
- Eble, C. 1996. *Slang & sociability: In-group language among college students*. The University of North Carolina Press.
- Facebook. 2010. Statement of rights and responsibilities. <http://www.facebook.com/terms.php>.
- Fehr, E., and Gächter, S. 2000. Cooperation and punishment in public goods experiments. *The Am. Econ. Rev.* 90(4):980–994.
- Gómez, V.; Kaltenbrunner, A.; and López, V. 2008. Statistical analysis of the social network and discussion threads in slashdot. In *Proc. of WWW*, 645–654. New York, NY, USA: ACM.
- Grudin, J. 2002. Group dynamics and ubiquitous computing. *Commun. ACM* 45(12):74–78.
- Hiltz, S.; Johnson, K.; and Turoff, M. 1986. Experiments in group decision making communication process and outcome in face-to-face versus computerized conferences. *Human Communication Research* 13(2):225–252.
- Hudson, J. M., and Bruckman, A. 2004. "Go away": Participant objections to being studied and the ethics of chatroom research. *The Information Society: An International Journal* 20(2):127.
- Jessup, L.; Connolly, T.; and Galegher, J. 1990. The effects of anonymity on GDSS group process with an idea-generating task. *MIS Quarterly* 14(3):313–321.
- Joachims, T.; Granka, L.; Pan, B.; Hembrooke, H.; and Gay, G. 2005. Accurately interpreting clickthrough data as implicit feedback. In *Proc. of SIGIR*, 154–161.
- Joyce, E., and Kraut, R. 2006. Predicting continued participation in newsgroups. *J. of Computer-Mediated Comm.* 11(3):723–747.
- Kirkpatrick, D. 2010. *The Facebook Effect*. Simon & Schuster.
- Lampe, C., and Resnick, P. 2004. Slash(dot) and burn: distributed moderation in a large online conversation space. In *Proc. of CHI*, 543–550.
- Leckart, S. 2009. The Official Prankonomy: From rickrolls to malware, a spectrum of stunts. *Wired* 17(9):91–93.
- Mackey, R. 2010. 'Operation Payback' Attacks Target MasterCard and PayPal Sites to Avenge WikiLeaks. *New York Times*.
- Mayer-Schonberger, V. 2009. *Delete: The Virtue of Forgetting in the Digital Age*. Princeton University Press.
- M.G. 2010. A Blizzard of protest over privacy. *Economist*.
- Millen, D. R., and Patterson, J. F. 2003. Identity disclosure and the creation of social capital. In *Proc. of CHI*, 720–721.
- Millen, D. R. 2000. Community portals and collective goods: Conversation archives as an information resource. *HICSS*.
- Naaman, M.; Boase, J.; and Lai, C. 2010. Is it really about me?: message content in social awareness streams. In *Proc. of CSCW*, 189–192.
- Poole, C. 2010. The case for anonymity online. *TED2010*.
- Rains, S. 2007. The impact of anonymity on perceptions of source credibility and influence in computer-mediated group communication: A test of two competing hypotheses. *Communication Research* 34(1):100.
- Ren, Y.; Kraut, R. E.; and Kiesler, S. forthcoming. Encouraging commitment in online communities. In Kraut, R., and Resnick, P., eds., *Evidence-based social design: Mining the social sciences to build online communities*. MIT Press.
- Rosen, J. 2010. The web means the end of forgetting. *The New York Times*.
- Rutkoff, A. 2007. With 'LOLcats' Internet Fad, Anyone Can Get In on the Joke. *Wall Street Journal* 25.
- Short, J.; Williams, E.; and Christie, B. 1976. *The social psychology of telecommunications*. Wiley London.
- Shuman, P. 2007. Fox 11 investigates: 'anonymous'. <http://www.youtube.com/watch?v=DNO6G4ApJQY>.
- Simmel, G. 1957. Fashion. *Am. J. of Soc.* 62(6):541–558.
- Smith, M., and Harper, D. 1995. Animal signals: models and terminology. *Journal of Theoretical Biology* 177(3):305–311.
- Sorgatz, R. 2009. Macroanonymous is the new microfamous. <http://fimoculous.com/archive/post-5738.cfm>.
- Suler, J. 2005. The online disinhibition effect. *International Journal of Applied Psychoanalytic Studies* 2(2):184–188.
- Szabo, G., and Huberman, B. A. 2010. Predicting the popularity of online content. *Commun. ACM* 53:80–88.
- Tanis, M., and Postmes, T. 2007. Two faces of anonymity: Paradoxical effects of cues to identity in CMC. *Computers in Human Behavior* 23(2):955–970.
- Thompson, P. A., and Ahn, D. 1992. To be or not to be: An exploration of e-prime, copula deletion and flaming in electronic mail. *Et Cetera: A Review of General Semantics* 146–164.