

## How Do People Express Identity Online, and Why Is This Important for Online Interaction?

Who are we when we go online, and how do we express that to others? Much internet activity consists of presenting oneself before others in one fashion or another. As a result, how we describe ourselves online shapes online interaction in critical ways. It is one of the most fundamental design choices that we make in creating online sites. To start exploring these issues, we need a nuanced understanding of face-to-face interaction. Erving Goffman's classic 1959 book, *The Presentation of Self in Everyday Life*, is a good place to start. Goffman wrote, "When the individual presents himself before others, his performance will tend to incorporate and exemplify the officially accredited values of the society, more so, in fact, than does his behavior as a whole . . . The world, in truth, is a wedding" (Goffman 1959, 35–36).

### The Presentation of Self in Everyday Life

What really is a wedding? At a wedding, everyone dresses up and collectively enacts a ritual that is meant to convey information to the entire group – these two people are now a family. Everyone at the wedding is playing a role – like bride/groom, mother of the bride/groom, the bride/groom's

school friends, the officiant, etc. Each participant is supposed to dress and behave in a certain way. We learn these social norms from personal and cultural examples.

I was more nervous before my own wedding than before giving significant professional talks to large audiences. The role of "professor" is one I'm comfortable with. The role of "bride" was unfamiliar. We are all always playing a role, and we play different roles on a daily basis. On a typical day I might drive my teenage son to school, stop off at the allergist for my allergy shot, buy a cup of tea at a café, meet with a graduate student, teach a class, and have a conference call with a collaborator. In each of these situations, I present different sides of who I am. Amy the mom, Amy the allergy sufferer, Amy the tea drinker, Amy the advisor, Amy the teacher, and Amy the research collaborator are all a bit different. I present myself differently in each of these settings. If I swapped two of those performances – for example, if rather than approaching the café counter and addressing the next available staff member to order tea, instead I stood at the front of the café facing everyone and projecting my voice to get everyone's attention as I would at the start of class – the results would be comic and awkward. My behavior in each setting is different, and who I am in each setting is also different. Together, all these aspects make up who I am.

Students in my "Design of Online Communities" class sometimes fall into the trap of thinking of these different facets of ourselves as facades masking our "true selves." In fact, there is no "one true self." It doesn't exist. We are all always performing, and who we are at the core is a synthesis of all these aspects (Turkle 1995, 261).

Goffman has a wealth of insights into human interaction and presentation of the self. It's worth going through these in detail, because most or all of these have analogs in online interaction. The original Goffman text is worth reading, though I need to warn you, as I always warn my class, that it has sections that are offensive (sexist, racist, classist, etc.) to a more modern sensibility.

Goffman begins simply enough by saying that "When an individual enters the presence of others, they commonly seek to acquire information about him or to bring into play information about him already possessed ... Information about the individual helps to define the situation, enabling others to know in advance what he will expect of them and what they may expect of him" (Goffman 1959, 1).

The most basic thing we need to know in many situations is someone's official role. To get my allergy shot, I need to know who (of the dozen people present at the allergist's office) is the nurse giving shots today. I arrive at the office and sign in, and the nurse calls my name when they are ready. I can infer that the person who calls my name from the room for allergy shots is the shot nurse for today. Their location (the shot room) and their behavior (calling my name after I've signed in) convey that information. But if I am confused, I can support my conclusion by their clothing – nurses at the office wear medical scrubs and a badge with their name.

Have you ever confused someone's role? It has happened to me many times, especially in clothing stores. I might ask, "Do you know if this comes in other colors?" and my embarrassed interlocutor replies, "I'm sorry, I don't

work here." I've also been mistaken for the employee. For example, at a recent trip to a clothing store, after I picked a sweater in my size from a pile, I then straightened the pile. A fellow customer saw me straightening clothing, and reasonably assumed that I worked there. These kinds of mistaken interactions can be uncomfortable because, as Goffman notes, having a particular role entitles a person to be treated in a particular way. Mistaken interactions happen often in clothing stores because staff typically do not wear a special uniform, so, unlike with the nurse at the allergist, we are missing attire as a cue to a person's role.

According to Goffman, information we might want to know about someone includes their socio-economic status, their conception of themselves, their attitude toward you, their competence, and their trustworthiness. Goffman assumes we can easily infer someone's gender, age, and race because we can see them. Online, we may (or may not) want ways to infer those aspects of personal identity.

Carriers of that information about the individual in face-to-face settings include clues from a person's behavior and appearance. This includes physical characteristics, how they dress, and how they hold themselves. We also infer information from a person's likelihood of being in a setting – the person calling my name from the shot room is probably the shot nurse, and the person standing at the front of the room at the start of class is probably the teacher.

In face-to-face settings, we are always calculating: What role does everyone here have? Who am I speaking to? As we will see, how we represent identity online shapes how these processes unfold in online interaction.

## Expressions Given and Expressions Given Off

In each setting, we are always consciously and/or unconsciously trying to convey an impression to others. However, how people actually interpret our performance may be different from what we intend. In Goffman's terms, there are *impressions given* (what we intend to convey) and *impressions given off* (what we unintentionally convey) (Goffman 1959, 2). Suppose a host cooks dinner for a visitor, and asks how the food is. The visitor might reply "delicious!" – intending to be graciously appreciative. However, the host might "take note the rapidity with which the visitor lifted his fork or spoon to his mouth, using these signs to check the stated feelings of the eater" (Goffman 1959, 7). The chef is checking impressions given off to try to develop a deeper insight than the guest's polite response. Goffman notes that because we know that people try to control what impression they are making, as interpreters of others' behavior we divide someone's performance into more and less easily controlled aspects. The dinner guest's speech is more easily controlled than the enthusiasm with which they move their fork.

Goffman's story of a guest approaching a Shetland cottage conveys an intriguing interplay between a performer trying to convey a particular image and a recipient interpreting that performance:

When a neighbor dropped by to have a cup of tea, he would ordinarily wear at least a hint of an expectant warm smile as he passed through the door in the cottage. Since lack of physical obstructions outside the cottage

and lack of light within usually made it possible to observe the visitor unobserved as he approached the house, islanders sometimes took pleasure in watching the visitor drop whatever expression he was manifesting and replace it with a sociable one just before reaching the door. However, some visitors, in appreciating that this examination was occurring, would blindly adopt a social face a long distance from the house, thus ensuring the projection of a constant image. (Goffman 1959, 8)

Here we have a kind of arms race between the visitor trying to convey a particular impression (*impressions given*) and what the host actually infers (*impressions given off*).

People's performances can be sincere or cynical. I can try to convey to you that I am an excellent financial advisor because I actually am skilled in that field (sincere) or because I am actually a swindler and would like to convince you to trust me so I can take your money (cynical). To add another layer of complexity, in the case of the sincere performance, I may be correct that I am a skilled advisor or I may be deluded. My interlocuter uses cues to interpret both the sincerity of my performance and my likely actual competence.

Human interaction is fundamentally collaborative. The start of an interaction establishes important expectations on the part of participants. Goffman shares the example of a teacher's approach to the first day of class:

The first day I get a new class in, I let them know who's boss ... You've got to start off tough, and then you can ease up as you go along. If you start off easy-going, when you try to get tough, they'll just look at you and laugh. (Becker 1952, 459; cited in Goffman 1959, 12)

Based on those initial expectations, participants establish a *working consensus* that helps get them through the situation. Once we have established that I am the patient and you are the allergy shot nurse, then we have a set of routines or “scripts” that we follow that help us get through the rest of the interaction smoothly (Schank and Abelson 2013).

### Fronts and Roles

Goffman defines a “front” as “that part of the individual’s performance which regularly functions in a general and fixed fashion to define the situation for those who observe the performance” (Goffman 1959, 2). He writes, “as part of personal front we may include: insignias of office or rank; clothing; sex, age, and racial characteristics; size and looks; posture; speech patterns; facial expressions; bodily gestures; and the like” (Goffman 1959, 24). Goffman divides these into appearance and manner (attitudes toward the current situation). We expect appearance and manner to be consistent, and to be consistent with the situation. Different roles may use some of the same elements as part of their front. In Goffman’s example, both chimney sweeps and perfume clerks of his time wore white lab coats.

In addition to appearance and manner, the third main element of the front is the setting, “furniture, décor, physical layout, and other background items which supply the scenery and stage props” (Goffman 1959, 22) for human activity that takes place there. Architecture and interior design create contexts that set the tone for human activity.

In some cases, there is a gap between the front and the person’s actual role. For example, “A patient will see his nurse stop at the next bed and chat for a moment or two with the patient there. He doesn’t know that she is observing the shallowness of his breathing and color and tone of his skin. He thinks she is just visiting” (Goffman 1959, 31). As a result, nurses sometimes don’t get the respect they deserve as skilled professionals. Similarly, much of the work of an undertaker is not visible to the customer, the bereaved family. Consequently, undertakers charge a great deal of money for the casket (something the family can see), and this helps cover the cost of all their services (Goffman 1959, 32).

### Identity Online: Usernames

Whether you are playing an official role in an online interaction or just hanging out, you have a personal front. People adapt how they interact with you based on how you present yourself – the elements of your personal front that you use.

At the simplest level, many online sites have textual usernames. In some cases, your username is chosen for you. Everyone at my son’s school has to use the username that is their last name, a period, and then their first name, so I am Bruckman.Amy. Standardizing usernames makes it easy to find anyone, and many information technology departments enforce this at schools and corporations.

In the 1990s through early 2000s, some sites for kids like Cartoonnetwork.com standardized usernames by letting you pick from a fixed list of adjectives and nouns

and adding a number – so you'd end up being something like BraveParrot331. The intent was to give people some degree of choice while making absolutely sure there would be no bad words included, and without having to pay a customer service representative to manage usernames. However, this style of username didn't work particularly well because the restricted choice didn't make you feel like this "was you."

In most cases on the internet, people get to choose their username, and you can tell a surprising amount about a person based on what name they select. Consider the following three usernames from the College of Computing at Georgia Tech:

jalisa

jkh

wedge

To protect people's privacy, I've chosen usernames from people no longer at Georgia Tech, but they were once active names. One of them is a professor, one is an undergraduate, and one is an administrative staff member. Can you guess who is who?

When I do this exercise in class, students almost always get it right. As you might guess, "wedge" is the undergraduate. Later in his career at Georgia Tech, when he became a master's student, he regretted the silly name and begged our technology services organization to change it to his first initial and last name. At the time it was against policy to allow username changes, but someone took pity on wedge and updated it. His sensibilities changed from the

time he was a seventeen-year-old freshman to the time he was a master's student trying to build a professional image.

User "jalisa" was an administrative staff member who conveyed a warm and friendly, informal tone by using her first name. User "jkh" was faculty, and used initials for a more professional self-presentation. How we choose to present ourselves online conveys how we see ourselves, and how we feel about the particular online setting.

One question that comes up for administrators is whether to allow obscene usernames. The answer depends on what kind of site you are running. On the bulletin board system ECHO (a kind of New York version of The WELL), founder Stacy Horn decided that she would "let people have whatever id they want. Go ahead, call yourself bigdick. See if I care. I think people would rather know than not know that you are the kind of person who would call themselves bigdick. It says something about you" (Horn 1998, 18). In Stacy's view, someone choosing a rude username is performing a public service by warning everyone what kind of person they are. Allowing rude usernames was a great strategy for ECHO, but wouldn't be appropriate for another site like, for example, one aimed at kids or affiliated with a brand.

Many sites require each username to be unique. Enough of the namespace is taken on large sites like Reddit that it can be hard to find a username that expresses what you want, unless you add a number to the end.

For email, the site hosting your account also conveys information about who you are. Posting from a university or corporate account conveys an official capacity.

In the 1990s, people with university accounts often looked down on people posting from large commercial providers like America Online (Donath 2002). This is less true today, as free email services like Google's Gmail have become widely accepted.

### Elements of Online Identity

Usernames are just one of myriad elements of personal identity we express on online sites. A host of design features contribute to people's online self-presentation. Elements of personal identity that often appear in the official part of an online profile include:

- gender
- topic-specific information (like a model motorcycle on a motorcycle forum, or a diagnosis on a medical forum)
- real name
- a profile picture/image (perhaps a photo, or an image the person chooses)
- a 3D avatar
- the person's past posting history on the site
- links to the person's presence on other sites
- profile text written by the person
- a "character class" the user has chosen (for games)
- membership in subgroups of the site (like guilds, for multiplayer games)
- achievements within the community (like Reddit karma)

Sites like dating sites may delve much deeper into personal information, with details like religion, political orientation,

sexual orientation, race, and age. On non-dating sites, race and age are rarely explicitly marked.

*When a designer chooses which elements of personal identity are easy to express on a site, the designer is essentially engineering a new personal front.* Imagine you were starting a face-to-face business, like installing solar panels. Decisions you might make include: Do your employees wear a uniform? If there is no explicit uniform, do you give them guidelines on what to wear? Do they have business cards? Clipboards? Do they drive a car with a company logo? All these decisions affect the impression your employees make on customers – you are designing their front. Similarly, when you are creating a new online site, you decide how users may present themselves to one another – should we have profile pictures? Should users be pseudonymous or identified by real names that we verify? Should we have people state their gender? Which elements you emphasize shapes the kinds of interactions that ensue.

### Gender Online

In 1983, Lindsy Van Gelder was on CompuServe, an early bulletin board system (BBS), and became friends with another user, Joan. Van Gelder tells Joan's story in her 1985 *Ms. Magazine* article, "The Strange Case of the Electronic Lover" (Van Gelder 1985). Joan was a neuropsychologist who had been severely injured in a car accident involving a drunk driver, and had ongoing challenges with both her mobility and speech. Talking on the BBS was her main social outlet. Online, Joan was outgoing, charming,

and generous with her time and money. She was also sexually aggressive. She identified as bisexual and was relentless in asking women on the board to have text-based sexual exchanges with her.

Joan introduced a friend on the BBS, Jane, to her friend Alex, a psychiatrist living in Manhattan. Jane and Alex hit it off, and Alex paid to fly Jane to visit him for a lavish weekend. They began a relationship.

By now you may have guessed the punchline to the story – “Joan” was really Alex. Alex had started the “Joan” account out of professional interest in what it felt like to be a woman, but then the experiment spiraled out of control and became something more elaborate and ultimately sinister.

Van Gelder’s story is timeless and raises a host of fascinating questions. First, she wonders “why a man has to put on electronic drag to experience intimacy, trust, and sharing” (Van Gelder 1985). Alex genuinely felt that people treated him differently when he was Joan. In what ways does gender shape how you treat people, online and in person? It’s well documented that accounts that present as female online often receive unwelcome romantic attention, ranging from flirting to harassment. Is there a degree of “intimacy, trust, and sharing” among women that men don’t experience? Alex’s experiment gives us a unique way to ask those questions and reflect.

Van Gelder writes that interaction online is “dizzily egalitarian, since the most important thing about oneself isn’t age, appearance, career success, health, race, gender, sexual preference, accent, or any of the other categories by which we normally judge each other, but one’s

mind” (Van Gelder 1985). It’s a beautiful vision, but to what degree is it true?

My observation is that for text-based communication, we still judge one another – but by a different set of markers. Notably, in writing, one judges by writing skill. Writing skill is a proxy for level of education and privilege. In more visual modes of communication, such as Snapchat and Instagram, traditional identity markers like age and race are back in the forefront of interaction.

Stacy Horn has a different view from Van Gelder. When Horn encounters people online who are gender-swapping, she feels that “You can’t always tell. At first. But you can often tell over time. The illusion of free and unbiased communication can only be maintained and then only briefly, as long as people hide. It’s a trick. In time, if you act like yourself, gender is revealed, because we do take our bodies with us. I don’t log on and suddenly forget I’m female. Oh, I’m online! Now I can forget a lifetime of socialization” (Horn 1998, 85). André Brock agrees, writing that “the digital is the mediator of embodiment and identity, not an escape from it” (Brock 2020, 20).

Questions about how one’s real-life gender shapes behavior patterns online became particularly challenging for Horn when she faced a dilemma: whether to let ECHO member “Embraceable Ewe” (who was in the process of transitioning from male to female) into the women-only forum on ECHO, WIT. Horn writes, “I’m in over my head... Is gender a biological or social construct?... If I let her into WIT, will it feel like there is a man in the room, or a woman?” (Horn 1998, 82). When Horn started ECHO,

she didn't anticipate having to be the arbiter of questions like the fundamental nature of gender!

To answer the question of whether to invite a transitioning person into an all-female group, we need to have a more nuanced understanding of why we are creating a single-gender space in the first place. Oldenburg notes that the Third Place (see Chapter 1) is often single-gender (Oldenburg 1989). What are we trying to achieve with explicitly single-gender online spaces? Today, many spaces self-select as predominantly one gender, but few have explicit gender restrictions. Ultimately, Horn allowed Embraceable to join WIT after she met a series of criteria, but writes that she regrets not simply accepting Embraceable as who she says she is. Our sensibilities on this topic have evolved since the 1990s.

Who is right, Van Gelder or Horn? Is the online world “dizzily egalitarian” or can you “always tell”? To explore this question, then-Georgia Tech graduate student Joshua Berman and I created a multiplayer identity game called The Turing Game, which we launched in 1999 (Berman and Bruckman 2001). The Turing Test is a challenge to see how far artificial intelligence has progressed – can a person tell the difference between written answers to questions from a person and a computer? In Alan Turing’s original paper about the Turing Test, he explains it first in terms of a gender test. Imagine that you have a man in one room and a woman in the other, and you can only communicate with them by slipping typed questions under the door. Could you tell who is who based on how they answer? Turing goes on to ask the reader to now imagine that there is a human

behind one door and a computer behind the other. However, Berman and I returned to his first example about gender.

In The Turing Game, a panel of volunteers all pretend to be a particular identity. Any identity is possible, and game types were chosen by users. Suppose everyone is pretending to be women. The panelists are supposed to answer questions as if they are women, and each chooses a woman’s name as a temporary pseudonym. Some really are women, and some are not. The audience votes to say who they believe. A moderator chooses among questions suggested by the audience. When the moderator decides the game is over, everyone’s real identity is revealed and a discussion ensues about how the audience knew the truth or how they were fooled. Here’s an example from a game where people were pretending to be women:

**QUESTION:** Describe your last really bad haircut.

**PENNY:** I had it layered and I got a perm. Since my hair is wavy, it was Annie style.

**WENDY:** Sophomore year, decided to cut it really short, and I looked like a little boy. My boyfriend was very disturbed.

Most people correctly guess that “Penny” is really a woman and “Wendy” is not. The reasons are interesting to reflect on. At the simplest level, the real woman shows a deeper knowledge of women’s hairstyles. But more interestingly, the person pretending to be a woman is worried about what her boyfriend thinks rather than what she herself thinks. In my experience, women tend to worry more about what they themselves think.

The game quickly became popular, with 12,000 users from all seven continents over the course of a year. People played with all different sorts of identity categories in addition to gender – like guessing who is really over 30, or who is Canadian. It was particularly popular with people who were transitioning their gender, people under house arrest, and the staff at a research station in Antarctica (there's not much to do there over the long, dark winter!). The point of the game was to get people discussing these issues of identity, and to develop some deeper insights into how people act online and how your personal identity shapes those interactions. And it did – sort of.

Analyzing data from games played, we found that people often made decisions based on stereotypes. For example, someone might say “I knew player two was a woman because she used such long sentences.” This is grounded in the idea that women use long, flowery sentences with lots of dependent clauses and men say things like “I'll be back.” The only problem is that this idea is just a stereotype. Studies of how many words men and women use per turn have different answers depending on whether they're talking to men or women, the medium of conversation, etc. A blanket “women talk more” stereotype is false (Baron 2004; Fox et al. 2007). The game surfaced stereotypes (some false, some true), and reinforced them, but didn't help people to distinguish stereotype from reality.

The second problem with The Turing Game as an interactive, philosophical experience is that the first thing you really want to know before judging someone's performance is how many times they have played the game. For

example, in one game in which people were pretending to be women, one question was: “What's your favorite alcoholic drink?” One respondent said, “scotch on the rocks.” The audience correctly reasoned that a man pretending to be a woman wouldn't say that, and that was true in that particular case. But now that you know that trick, you might consider trying a double reversal to fool people (like the movie *Victor Victoria*).

The ultimate lesson from The Turing Game is that we need to work hard to distinguish between stereotypes and empirically verifiable reality. And those are moving targets – both the stereotypes and the reality are always evolving.

In the end, Van Gelder and Horn are both partially right. Sometimes this new mode of interaction can be surprisingly egalitarian, and sometimes not. The exact details of how sites are designed and how people can present themselves matter.

### Identity Workshop

“Joan” was playing with aspects of identity online. Online, you can pretend to be someone you're not, and that process can help you reflect on your face-to-face identity and sense of self. The internet can be a kind of “identity workshop.” That was the title of my final paper in Sherry Turkle's class on sociology of technology in 1992. Taking Turkle's class and serving as her research assistant for part of her book *Life on the Screen* (Turkle 1995) were transformative experiences in my graduate-school career.

Turkle thinks deeply about the psychological foundations of our interactions with computers. In her 1984 book, *The Second Self, Computers and the Human Spirit*, she wrote about Deborah, a young girl who was concerned about issues of control in her life. When Deborah learned to program a computer in Logo, she was engaging in an intellectual exploration of issues of control, a key emotional and psychological issue she was struggling with at the time (Turkle 1984). Most research on humans and computers stays relatively at the surface: this is what people like, this is what people do. Turkle pushes deeper to understand *why* people like certain activities, and why they engage in them. And the answers to those questions are not superficial – they are about who we are and who we wish to become in the most profound sense.

My class paper explored concepts of identity in MUDs, text-based virtual-reality environments that were popular at the time (see Chapter 1). A MUD is like any massively multiplayer game, except that the whole world is created in text (for example, the description of The Living Room on a MUD called LambdaMOO begins: “It is very bright, open, and airy here, with large plate-glass windows looking southward over the pool to the gardens beyond . . .”). In addition to describing objects and places, you can describe yourself. Because you can easily describe yourself as anything you want, it becomes easy to experiment with identity. The following summer, I helped Turkle study MUDs further, assisting with research for *Life on the Screen* (Turkle 1995).

Turkle is a psychoanalyst, and one fundamental question she asked about this online identity play was:

Does it help? If people are playing with concepts of identity, are they growing in a healthy way as a result? We wrestled with this basic question all summer, and had no conclusion by September when I went back to my regular research at the Media Lab. Her final conclusion in the book is: sometimes. For some of the people we interviewed, they explored personal issues they were struggling with for a period of time, and later made life changes that showed growth. Others just struggled more and more. In psychological terms, some were “working through” and others were “acting out.” The technology is evocative, but a range of things can happen as a result.

### Identity Deception

Gender-swapping is just one form of identity deception that takes place online. People can present themselves differently from their real selves in any dimension of personal identity. In some contexts, your real identity is expected, and deviating from that is breaking a social norm. In others, identity play is welcome and expected.

Understanding deception is one way to develop insights into how identity functions in online social groups more generally. Judith Donath explains deception in terms of ideas borrowed from research on animal behavior: *assessment signals* and *conventional signals* (Donath 2002). An assessment signal is a reliable signal that demonstrates the trait being shown. For example, having big muscles is an assessment signal for strength. Displaying an assessment signal is costly to the sender, because they need to really

have that trait. On the other hand, a conventional signal is one that's easy to fake. For example, wearing a "Gold's Gym" T-shirt is a conventional signal for strength. When we interact online, some signals become easier to fake. It takes quite a bit of work to present as someone of a different gender in a face-to-face setting (with clothing, changed personal manner, etc.), but online it may be as simple as one click. Donath notes that the world of deceptive and true signals creates an ecosystem. In butterflies, there are monarch butterflies that use their bright coloring to tell birds that they are not tasty to eat. There are also moths that mimic monarchs in appearance. The moths are tasty, but birds avoid them because they look like monarchs. The more moths imitate monarchs, the more a bird might consider eating something that looks like a monarch – the odds of it being tasty go up. Having lots of deceivers brings a cost to the non-deceptive members of the system. Many online interactions are low-stakes – it may not really matter if the person you are chatting with is not really as they describe themselves (as long as you're not relying on them for medical or legal advice!). One higher-stakes example is dating sites. If someone lies about who they are, people may waste time they could spend talking with someone more suitable, and may have uncomfortable face-to-face encounters. Toma et al. wondered how much people lie on dating sites. To find out, they had people on dating sites come into a lab setting and compared their online profiles to their real selves. They looked at people's driver's licenses to confirm age, measured their height, and

had them step on a scale to get an accurate weight. They found that people generally do lie – but just a little bit. If you are hoping to eventually meet someone face-to-face, lying a lot isn't strategic. However, stretching the truth just a bit is so tempting that most people do it. Eight out of ten people lied in at least one category. They write, "weight was the most frequently lied about attribute, followed by height, and least of all age. For those identified as lying on an item, the magnitude of the deception was usually small. The average deception for height was only 2.09% of the participants' actual height, 5.5% of the participants' actual weight, and 1.4% of the participants' actual age" (Toma et al. 2008). The ecosystem of people being more or less truthful on online dating sites is like the ecosystem of butterflies and moths – deceptive presentations create a cost for others.

### Age and Race

Gender is the most studied component of online identity, but is not necessarily the most important. It's an interesting research question in itself to think about why researchers to date have been so focused on gender and have paid less attention to other aspects of identity. Other factors like age arguably shape how you interact with others in face-to-face interactions more.

Age is too often neglected in research. In her master's thesis on online role-playing of Harry Potter fan fiction on LiveJournal, Casey Fiesler found that many people were role-playing erotic scenes, not realizing that their scene partner might be under-age. Acting out a sex

scene with a teenager is not only creepy, it's potentially a serious illegal act. It just never occurred to the LiveJournal role-players to check someone's age, especially since the social norm of the group was to not ask people about their real-life identities (Fiesler 2007).

Age is particularly salient as an aspect of identity in countries where age strongly shapes interpersonal interaction. In Korea, the language itself changes depending on how old the person you are addressing is – you use different words to talk to someone older than to talk to a peer. A Korean student in my “Design of Online Communities” class once told me that he believes that the notable popularity of massively multiplayer online games in Korea is linked to that fact. Online, Koreans address everyone as a peer, and he believes that makes online interaction particularly satisfying for them. It would be interesting to try to study whether the student’s intuition is correct.

Race is often not expressed in online interaction. André Brock laments that the default internet user is middle class, white, and heterosexual. As we’ll see, spaces like Black Twitter help to decentre that whiteness, making a space where Black people “do not feel compelled to hide or change their cultural particularities” (Brock 2020, 87).

### **Identity and Communities Focused on a Specific Demographic**

The internet was initially developed from the 1960s to early 1980s by largely white, male engineers at research universities. In its formative years, that demographic was dominant.

Personal identity often went unmarked, and if it was unmarked people assumed others they were interacting with were white (or Asian) males – the stereotypes for computer scientists and engineers at the time. As the internet became a mass phenomenon in the 1990s, this changed. During the first dotcom boom in the late 1990s, when early internet companies thrived and venture capital was widely available for any internet idea, a number of niche online sites were created to appeal to other demographics. For example, the site iVillage was launched in 1995 to appeal to women. Around the same time, Third Age was launched to appeal to older adults, and Black Planet to appeal to African-Americans.

In the early days of iVillage, developers tried to create content to appeal to people like themselves – technologically literate, college-educated women. Over time as actual internet users became more diverse, the site was re-focused on the kind of content that you might find in a women’s magazine sold in a supermarket checkout line, with articles about topics like weight loss, dating, clothes, and makeup. A similar pattern occurred on many sites. When I interned at Third Age, a website for older adults, in summer 1997, the site was investing lots of money in hiring freelance writers to write thoughtful long-form articles on topics like how your worldview changes when you retire, and how to manage finances on a fixed income. However, looking at the web server logs, they realized that those articles were hardly being read. Instead, users were doing puzzles and using the dating personals. The same lesson again can be seen when Christian Sandvig studied a community network installed on a Native American reservation in 2004. Sandvig documents how tribe members emphasize the

the educational value of the network when describing it to visitors, but in day-to-day practice their use looks like the use of any other group of people – lots of gaming and online shopping (Sandvig 2012). The lesson here is a deeply human one: people are people. If people in general are interested in a particular kind of content, people in a subgroup will likely have similar interests.

I draw two critical lessons from these cases. First, to build a successful site, it's important to do user-centered design. Online community designers need to gather actual data from potential users before assuming what those people will want. Second, it's important to have online sites that represent a wide range of interests for different people – both ones with makeup tips and ones with essays on the future of intersectional feminism. However, as a site designer, if you are driven solely by maximizing your page views and hence advertising revenue, then you'll get all dating tips and no thoughtful essays. *Being driven solely by the profit motive is breaking the internet.* Yes, to create the highest possible membership on a women's site you need to post stereotypical content. However, it is possible to deliberately choose to promote a different sort of content, if maximizing advertising revenue is not your only concern.

As an example, the website Rookie Magazine ([www.rookiemag.com](http://www.rookiemag.com)) was created by a fifteen-year-old girl and had articles written by girls and young women on topics like "Meaningful Transformations, my favorite movies about change" and "The Last Train Ride, a story about friendships that fade." It was deliberately non-commercial, and focused on content that matched the founder's values.

Rookie Magazine published issues for seven years, and then folded in 2018 due to business issues. It's interesting to think about whether a site like this could be financially sustainable. One thing, though, is clear: It is impossible if initial funding comes from venture capitalists who demand a high rate of return on their investment.

### **Subgroup Conversation in a Public Space: Black Twitter**

Most communities that focus on one particular demographic try to create a separate space just for that group. Membership may be explicitly restricted, or expectations of who belongs may lead to self-selection. For example, some Facebook groups for mothers also allow fathers to participate, and others don't.

An intriguing exception to this pattern of membership is Black Twitter (Brock 2012; Manjoo 2019). André Brock writes that "Black Twitter is an online gathering (not quite a community) of Twitter users who identify as Black and employ Twitter features to perform Black discourses, share Black cultural commonplaces, and build social affinities" (Brock 2020, 81). While sites like Reddit divide people into sub-communities and let social norms evolve differently in each subgroup, Twitter is one huge, undivided stream of content. Your main Twitter feed shows content only from people you follow. You can also search for terms or hashtags (#term), and interact with others who post about those topics. Conversations of interest to Black members of Twitter emerge from relevant hashtags, and also from follower relationships.

Black Twitter builds on African-American cultural practices. For example, Sarah Florini discusses the practice of “signifyin’,” “a genre of linguistic performance that allows for communication of multiple levels of meaning simultaneously, most frequently involving wordplay and misdirection” (Florini 2014, 224). Florini writes that “Signifyin’ serves as an interactional framework that allows Black Twitter users to align themselves with Black oral traditions, to index Black cultural practices, to enact Black subjectivities, and to communicate shared knowledge and experiences” (Florini 2014, 224). Similarly, it builds on oral traditions of call-and-response (Manjoo 2019). Some of the discourse is playful and focused on humor, creating a supportive sense of being in the presence of like-minded others. Black Twitter also has an activist focus, using the platform to draw attention to social issues like police brutality (De Choudhury et al. 2016).

While a site like a subreddit or a Facebook group has a space that is dedicated to that topic and group of members, Twitter conversation is all public. A small fraction of users have private accounts that only approved people can see; however, this is uncommon. As a result, this in-crowd conversation is taking place in view of a diverse audience. This at times is an opportunity for mutual understanding across groups, but also a catalyst for tension among those groups.

### Privacy

We don’t always want to reveal everything about ourselves online. Before I discuss online anonymity and

pseudonymity, I need to take a detour to introduce basic ideas about privacy. What is privacy and why does it matter?

Privacy is the right to control a “zone of accessibility” around yourself. It includes freedom from intrusion, the ability to control information about oneself, and freedom from surveillance (Baase 2013, 48). Privacy lets us be ourselves, lets us remove our “public persona,” and is arguably important for individuality and freedom (Quinn 2017). Simson Garfinkel writes that:

The problem with the world “privacy” is that it falls short of conveying the really big picture. Privacy isn’t just about hiding things. It’s about self-possession, autonomy, and integrity. As we move into the computerized world of the twenty-first century, privacy will be one of our most important civil rights. But this right of privacy isn’t the right of people to close their doors and pull down their window shades – perhaps because they want to engage in some sort of illicit or illegal activity. It’s the right of people to control what details about their lives stay inside their own house and what leaks to outside. (Garfinkel 2000, 4)

I teach about privacy each year in my class “Computing, Society, and Professionalism” (our required undergraduate class on ethics and social implications of technology). Over the years, I’ve found that the most important concept is the trade-off between *free-market approaches* and *consumer-protection approaches* (Baase 2013, 107–110).

In a *free-market approach*, companies may treat user privacy as they wish. Consumers are encouraged to use voice and exit (see Chapter 6) to respond if they are

not satisfied with how a company is treating them. In other words, if you don't like the privacy protections on a site, you can speak up or go to another site.

On the other hand, a *consumer-protection approach* suggests that people are often not able to make well-informed decisions about their privacy. First, they may not even know what data is being collected about them. This is particularly true of "invisible" forms of data collection, like tracking RFID tags or GPS coordinates. Second, most people don't read privacy policies on sites they use. They don't and they couldn't if they tried. The reading level of these policies is well above the level of the average internet user (Jensen and Potts 2004), and reading all the policies for sites you use would simply take too much time (McDonald and Cranor 2008). People are not in a good position to make smart choices about their own privacy.

MIT professor Jerry Saltzer once told me, "privacy is a database correlation problem." What ends up being most invasive is not the data from any single site, but the cross-referencing of data across different sites. Also problematic is the use of data for secondary purposes – uses other than the ones the data was collected for. If people have no reasonable way to envision secondary uses of data, they can't make informed choices about how much data to share.

If people can't make informed choices about how much personal information to share, this argues that we need to pass laws to protect individuals' privacy. However, some argue that the more restrictions we put on how businesses deal with user data, the more we reduce productivity and slow innovation. There are advantages and

disadvantages to each approach. It is a trade-off – one that we need to carefully assess in each situation.

### Social Networks and Privacy

My second post when I started my blog in January 2010 is entitled "Amy's Prediction: In 20 Years No One Will Be Qualified to Be President":

Today's teens are pouring their most personal thoughts onto the Internet. They flirt, they gossip, they angst, they brag about being naughty—just like we did when we were teens. Except the problem is, the Internet is a surprisingly persistent medium.

An old joke says that taking information off the Internet is like taking pee out of a pool. Sure you deleted it, but did the server keep a backup? There's likely a backup on Brewster Kahle's Internet Archive, <http://archive.org>. Before you decided to delete it, did a friend save a copy? When you post information online, you lose control of it.

Teens say the most amazing things. My friends and I had a great deal of fun, and I'm relieved to say it's all forgotten or at least not documented in my own words or photos. (If I appear doing anything unseemly in my friend Anne Mini's new novel, I can simply deny it!) If all of our coming-of age angst was saved for posterity, I'd be appalled. I think most people look back on their teen and young adult years that way. At least I hope they do.

What happens when young adult antics are archived? The thought gives one pause. Will the bride data mine the groom before the wedding (or vice versa)? Will the colleague with an axe to grind dig up ancient history to

use as a weapon? Are we entering a new age of harassment by ancient history, a golden age of blackmail?

I suspect that most teen and young adult antics will stay obscure, and if they're uncovered folks will mostly just laugh and reminisce. But there's one special category of people who may not get away so easy: public figures. Actors, musicians, and athletes can probably survive the scrutiny. But what about politicians? We still elected Bill Clinton, because he said he "didn't inhale." What happens when the future political candidate is inhaling on camera, memorialized for posterity?

I see a few possible outcomes. One is that teens over time will learn to be more careful with their personal information. This I think is inevitable. Which leads us to the prospect that we will have one lost generation of potential future politicians—the generation who didn't yet know to be careful about their personal information online. Like the donut hole in Medicare coverage, we'll have a lost zone between those too old to have been online much and those young enough to know to be at least a bit careful.

Another potential outcome is that we as a culture will learn to be more tolerant of what people do in their personal lives, especially as youth. Europeans tend to be somewhat more tolerant already—to draw a clearer line between personal and professional behavior. Americans are plagued by an endearing notion of "Character"—that what we do in our personal lives speaks to our fitness for professional tasks. When complete lives are increasingly archived, we may need to step back from that ideal and let our leaders be human. (Bruckman 2010)

Eleven years later, I'd say my prediction is on track. As we interact online, we continually leak little details about

ourselves. The problem is usually not any one piece of information, but the total picture that comes from the synthesis of those details.

Many years ago in my "Online Communities" class (around 2001), master's student Alisa Bandlow studied people role-playing on the site LiveJournal. Some of the content included erotica. Bandlow asked one of her interview subjects how they felt about the fact that all this content was visible online, linked to their account which used their real name. After the interview, the subject *deleted their account and all their postings*. It hadn't previously crossed their mind to be concerned. Over time, people are becoming more aware that they need to be careful about their online privacy. One way to protect your privacy is to not use your real name for segments of your online activity.

### From Anonymous to Identified

I talked earlier about email account names, and the ways that the account name you choose or are assigned expresses who you are to others. This holds not just for email addresses, but for account names we sign up for on online platforms.

An online username is a "signifier." It's a symbol that points to a person in the real world, the "signified." Semiotics studies the relationship between signifiers and signifieds, and how language makes meaning (de Saussure 1959). A signifier is typically arbitrary. There's nothing about the name "Amy" that is specific to me—it's just a sequence of sounds.

Many online usernames are "identified," which in semiotic terms means there is a publicly visible relationship

between the online account (the signifier) and the real-world entity (the person). Other sites can be “anonymous” or “pseudonymous.” These are sometimes confused and the distinction is important. On a pseudonymous site, a user has a persistent pseudonym that refers to them. Reddit is an example. For example, “shiruken” is the pseudonym of one of my fellow Reddit moderators. I know that he is male, that he is a biomedical engineer, that he runs our statistics on who had the most mod actions each month, and that he has a variety of scientific interests. I don’t actually know his real name or where he lives. In semiotic terms, there is still a relationship here between the signifier and the signified. His Reddit username is not his real name, but it is still a signifier that refers to him, the real person, as the signified.

Pseudonyms operate similarly to “real names.” The name I was given at birth is a signifier that has a legal relationship to me, the human. Just like pseudonyms, the relationship between this signifier and the signified can be slippery. I might legally change my name, use a fake name, or be a victim of mistaken identity. For both real names and pseudonyms, there is a signifier with a complicated relationship to the signified. My “real” name has a legal status but otherwise operates in similar ways to pseudonyms I use. What is a “real name” can be surprisingly controversial. Facebook requires users to use their real name, and users can report another account for violating the real names policy. Facebook reviews millions of such reports per year. Enforcement of this policy has caused problems for a variety of groups, including people transitioning their gender, drag queens who wish to use a stage name, Native

Americans who wish to use a tribal name, and survivors of domestic abuse, bullying, and stalking.<sup>1</sup> Enforcing a particular legal notion of a real name is a form of oppression in these cases (boyd 2011). In 2015, Facebook introduced explicit support for name changes in those special cases (Facebook 2015), but there continues to be tension around this issue.

“Identified,” “pseudonymous,” and “anonymous” are not discrete categories, but rather a continuum. As we move from more nearly identified to more nearly anonymous, the relationship between the signifier and the signified gets looser, but it’s still there.

When you use a pseudonym over a long period of time, you continually leak little details about yourself (unless you’re careful not to). In the course of conversation, you might mention things you like and dislike, places you’ve been, schools you’ve attended, and more. The longer you use a pseudonym, the more it de facto identifies the real person. Each leaked detail links the signifier and signified more tightly together. For this reason, users on Reddit will often create a temporary (“throwaway”) account when they really want privacy, often using “throwaway” in the name. If you only use an account for one discussion (usually on a difficult or embarrassing topic), then it is less tied to your real identity than an account that has an accumulation of leaked details about you (Leavitt 2015). Ammari et al. found that Reddit users on parenting boards who used throwaway accounts

were more likely to discuss stigmatized topics like pregnancy loss, divorce, or custody issues. Posts by throwaway accounts were more likely to get a response, and got more long responses (Ammari et al. 2019).

While a site like Reddit operates with pseudonyms, sites like 4chan are closer to truly anonymous. Posting on 4chan does not require any kind of account or login. Furthermore, most posts disappear within minutes and are not archived (Bernstein et al. 2011). Even if you leak something about what you are interested in with a particular post, it is not linked to all your other posts.

Even at the ends of the spectrum from most nearly identified to most nearly anonymous, there is still a degree of uncertainty. You can imagine that a determined person (e.g., a spy) could falsify an identified account, pretending to be someone they are not. Further, you can imagine that even on an anonymous site, if a government had a compelling need to find someone, government agencies might be able to trace internet traffic to help locate the real person. As a result, it's not a matter of whether we are identified or anonymous – it's more a matter of how hard it is to trace the link from the online presence to the real person. This ranges from easy to very hard.

### Why We Need Pseudonymity

Where to place a site on the continuum from anonymous to identified is a key design decision. Anonymity/pseudonymity is particularly valuable for situations where people may wish to disclose highly personal information – such as in

support groups. It's also critical for people who live under oppressive regimes.

While those dramatic cases are important, there are many reasons why ordinary people might want to express themselves pseudonymously on a day-to-day basis. First is the issue of “context collapse” (Vitak 2012). We all present different aspects of who we are in different contexts. For example, you may want to discuss politics on the internet, but not share your political views with your professional colleagues. A functioning democracy requires places to debate issues of the day, but a professional workplace is typically not an appropriate venue for those discussions. Similarly, you might not want to make your interests and hobbies part of a professional persona. The inverse is also true: If you are playing a massively multiplayer video game, you may not want to disclose that you are a professor for your day job! Pseudonymity lets me be both a serious scholar and a fan of the television show *Grey's Anatomy* and to keep those personas separate.

The second issue is permanence. I am eternally grateful that the internet did not exist when I was a teenager – I'm sure I would have said things that mortify me today. Does everything I say need to be part of my permanent record? Judith Donath writes that online reviews, the product review you generously provided for an underarm deodorant or for books about coping with binge eating or bed-wetting, will, if written under your real name, be part of your online portrait, what your neighbors, kids and random strangers see about you. Online, words persist forever, in vast searchable

databases. Anything you say or do using your real name is permanently attached to it. (Donath 2014)

As a culture, we have not yet quite sorted out how to handle the problem of permanence of information. If someone did something culturally insensitive or racist (like wearing “blackface”) decades ago, does that disqualify them for public office now? When is saying “sorry” enough, and when is it not? Because more people are becoming aware of the risks of permanence, sites like Snapchat that make communication ephemeral are becoming more popular. A picture shared on Snapchat goes away in a day, and the system does its best to prevent others from making a copy. (They unfortunately can’t stop people from using another camera to photograph the screen.) Snapchat is used by teenagers for hanging out, while a site with permanence like Instagram is used for curating a self-presentation.

### Anonymity and Accountability

Pseudonymity is not always appropriate. In 1995, I started a site called MediaMOO, a professional community for media researchers (Bruckman and Resnick 1995). A MOO is an end-user programmable, text-based virtual world. The MOO software was created by Pavel Curtis, then at Xerox PARC. MediaMOO was kind of like an endless conference reception for a conference on media and human-computer interaction. When I first created the site, I allowed people to have two accounts: one anonymous and one identified by a member’s real name. Curtis told me that was silly – why would anyone want to go to a professional conference and

hide their name badge or wear a bag over their head? The whole point was to meet other people and network. MediaMOO tried the two-account solution for a few years, but I eventually realized that Curtis was right. I changed our policy to specify that new accounts were to be identified only. Old anonymous ones were allowed to persist, but were rarely used. MediaMOO is an example of a context in which being identified is more appropriate.

Anonymity/pseudonymity also have a major downside: The lack of accountability can lead to bad behavior. The next chapter will talk in detail about bad behavior online and what to do about it. The image boards 4chan and 8chan are intriguing examples of sites that are more nearly truly anonymous. This has led to a proliferation of content that many people find extremely offensive, including highly explicit images, racism, anti-Semitism, and ideas from the alt-right. The sites often also originate attacks on others, such as “brigading,” where individuals encourage large numbers of people to harass individuals on other sites. In 2019, perpetrators of mass shootings posted about their plans on 8chan three separate times. As a result, 8chan’s domain name service was shut down, but it remains accessible on the dark web.

It’s important to note that 4chan is also a sort of “internet meme factory” which has given us things like “LOLcats,” and also held animal abusers accountable (Bernstein et al. 2011). Gabriella Coleman tells in detail the story of how the group “Anonymous” originated on 4chan, and over time parts of the hacker group evolved into a “hacktivist” group trying to work toward social justice (Coleman 2014).

It's possible to be well behaved on an anonymous site, and to create havoc on an identified site. The difference is that an agent of chaos on an identified site is more likely to experience real-world consequences, since the link between their online presence and real identity is stronger. Those consequences serve as a deterrent to bad behavior.

### The Future of Online Identity

The current mechanisms for representing identity online are woefully inadequate. Last week I was followed on Twitter by several fake accounts. I suspected they were fake because their interests had no relationship to mine, and they had few tweets to date. A quick image search showed that they had profile photos stolen from other online sources. Many such accounts are aimed at political manipulation – they create an online history that makes them appear to be a normal person, and then later use the account to spread political propaganda. There should be a way to simply say “that’s not a real person – they can’t follow me.” Similarly, there should be a way to say “that’s not a real company – they can’t send me email,” or “that’s a child – they shouldn’t be using this site with explicit sexual content.” All of these problems – political manipulation of elections, spam, children exposed to explicit content, and more – are fundamentally failures of mechanisms for online identity.

Would it be possible to design a new *identity layer* for the internet? I can make a better decision about whether to talk to a potential new contact if I know whether they really are affiliated with the company or school they claim,

whether they really are a friend of one of my friends, etc. If we could tell who is who, we could stop spam, stop deliberate manipulation of our public discourse by foreign operatives, prevent kids from getting to inappropriate content, prevent phishing attacks, and more.

Identity goes hand-in-hand with *reputation*. I imagine a system of *reputation servers*. A reputation server could verify that someone isn’t a known bad actor, and estimate their qualifications with regard to a particular topic. Reputation needs to be multi-dimensional – that is, I might have a strong reputation as an expert on human-computer interaction but a weak reputation as a keeper of tropical fish (since I have just a small freshwater tank with easy-to-care-for fish). Further, some aspects of reputation may need to be relative to an individual’s position on a particular issue – in other words, Jane has an excellent reputation as a chef for omnivores, but not for vegetarians. There would have to be competing reputation services, so individuals could pick a service they trust. Further, the design of such a system would have to make careful use of cryptography to protect people’s privacy. Unfortunately, it’s hard to imagine identity and reputation data that can’t be manipulated to mine data about citizens. Further, being identified creates serious risks for members of marginalized groups and people who live under oppressive regimes. How to improve our management of online identity without putting those groups at even more risk may be an impossibly hard design problem. But that doesn’t mean our current system is the best that is possible. The challenge is: Can we design

improved versions of identity and reputation that help us tell who is who, but in a privacy-preserving way? The potential risks and gains are both high.

### Theoretical Summary

As Erving Goffman described in *The Presentation of Self in Everyday Life* (Goffman 1959), we are all always performing for others. We try to make a particular kind of impression on people we interact with, and we do that differently in each social context. There is no “true self” – we are all a synthesis of different aspects of the self that we present.

To do this, we use “fronts.” A front includes both our appearance and manner. An observer divides people’s self-presentation into more and less easily controlled aspects, and judges performances more on the less easily controlled. The impression a performer wishes to make is their “impressions given,” and the impression the observers actually form are the “impressions given off.” Performances may be sincere or cynical.

Online, we similarly are always trying to make an impression. The key difference is that all the elements of our expressive vocabulary are explicitly chosen by site designers.

One critical choice in site design is whether user-names should be more nearly anonymous, pseudonymous, or identified. A name (whether legal real-world name or online account name) is a “signifier” that has a relationship to a “signified.” As we move along the spectrum from anonymous to identified, the relationship between the

signifier and signified gets stronger. Where a site should be along this continuum depends on the site’s purpose.

There are legitimate uses of anonymous communications, notably support groups and communication under oppressive regimes. However, being more nearly anonymous can lead to reduced accountability.

### Practical Implications

Identity is fundamental to the design of social media and the design of the internet. In person, we dress and behave differently in different social situations. Online, the same mechanisms operate. However, online site designers have the power to choose what elements of identity it is possible to express and which are emphasized. These choices profoundly shape online interaction.