Identity and Deception in the Virtual Community

Judith S. Donath
MIT Media Lab

Prepared for: Kollock, P. and Smith M. (eds). *Communities in Cyberspace*.
London: Routledge

Available in Croatian, as Identitet i prijevare u virtualnoj zajednici, (translation courtesy of The Institute of Ethnology and Folklore Research) or in Greek, as ΤΑΥΤΟΤΗΤΑ ΚΑΙ ΕΞΑΠΑΤΗΣΗ ΣΤΗΝ ΔΥΝΗΤΙΚΗ (VIRTUAL) ΚΟΙΝΟΤΗΤΑ

Identity and the virtual community

Identity plays a key role in virtual communities. In communication, which is the primary activity, knowing the identity of those with whom you communicate is essential for understanding and evaluating an interaction. Yet in the disembodied world of the virtual community, identity is also ambiguous. Many of the basic cues about personality and social role we are accustomed to in the physical world are absent. The goal of this paper is to understand how identity is established in an online community and to examine the effects of identity deception and the conditions that give rise to it.

In the physical world there is an inherent unity to the self, for the body provides a compelling and convenient definition of identity. The norm is: one body, one identity. Though the self may be complex and mutable over time and circumstance, the body provides a stabilizing anchor. Said Sartre in *Being and Nothingness*, ``I am my body to the extent that I am,'' The virtual world is different. It is composed of information rather than matter. Information spreads and diffuses; there is no law of the conservation of information. The inhabitants of this impalpable space are also diffuse, free from the body's unifying anchor. One can have, some claim, as many electronic personas as one has time and energy to create.

``One can have...?'' Who is this ``one''? It is, of course, the embodied self, the body that is synonymous with identity, the body at the keyboard. The two worlds are not really disjoint. While it is true that a single person can create multiple electronic identities that are linked only by their common progenitor, that link, though invisible in the virtual world, is of great significance. What is the relationship among multiple personas sharing a single progenitor? Do virtual personas inherit the qualities - and responsibilities - of their creators? Such questions bring a fresh approach to ancient inquiries into the relationship between the self and the body - and a fresh urgency. Online communities are growing rapidly and their participants face these questions, not as hypothetical thought experiments, but as basic issues in their daily existence. A man creates a female identity; a high school student claims to be an expert on viruses. Other explorers in virtual space develop relationships with the ostensible female, relationships based on deep-seated assumptions about gender and their own sexuality; patients desperate for a cure read the virtual virologist's pronouncements on new AIDS treatments, believing them to be backed by real-world knowledge. For assessing the reliability of information and the trustworthiness of a confidant, identity is essential. And care of one's own identity, one's reputation, is fundamental to the formation of community.

Identity cues are sparse in the virtual world, but not non-existent. People become attuned to the nuances of email addresses and signature styles. New phrases evolve that mark their users as members of a chosen subculture. Virtual reputations are established and impugned. By looking

closely at these cues, at how they work and when they fail, we can learn a great deal about how to build vibrant on-line environments.

The Usenet environment

This paper examines identity and deception in the context of the Usenet newsgroups. Although technically simple - they are essentially structured bulletin boards - a complex social structure has evolved within them. Unlike many MUDs, which are intended as fantasy worlds, most of Usenet is meant to be non-fiction; the basic premise is that the users are who they claim to be. There is, however, a significant variance between newsgroups as to what constitutes a real or legitimate identity. And there are numerous cases of identity deception, from the pseudo-naive trolls to the name-switching spammers.

People participate in Usenet newsgroups for a variety of reasons. They may be seeking information or companionship, advocating an operating system or a religion. As in the real world, their motivations may be complex: both the desire to be helpful and the desire to be noticed may prompt the writing of a lengthy exposition. For most participants, identity - both the establishment of their own reputation and the recognition of others - plays a vital role.

Information exchange is a basic function of Usenet. Requests for information are very common and answers, both right and wrong, are usually forthcoming. In the real world we may believe a story if it was published in The Wall Street Journal and dismiss it if it appeared in The National Enquirer. With Usenet, there is no editorial board ensuring standards of reliability; each posting comes direct from the writer. Thus, the writer's identity - in particular, claims of real-world expertise or history of accurate online contributions - plays an important role in judging the veracity of an article. Similarly, knowing the writer's motivation - e.g. political beliefs, professional affiliations, personal relationships - can greatly affect how we interpret his or her statements. Is the persuasive posting about the virtues of a new compiler coming from a programmer who has evaluated its code output or from a marketer of the product? The reader who knows that the author stands to gain from promoting a product or an idea is likely to doubt the veracity of the claims [Aronson 95].

The cost of identity deception to the information-seeking reader is potentially high. Misinformation, from poor nutritional advice to erroneous interpretations of drug-smuggling law, is easy to find on the net - and it is more likely to be believed when offered by one who is perceived to be an expert [Aronson 95]. The limited identity cues may make people accept at face value a writer's claims of credibility: it may take a long time - and a history of dubious postings - until people start to wonder about the actual knowledge on a self-proclaimed expert.

Providing affiliation and support is another important function of Usenet [Sproull & Kiesler 91, Wellman 97]. Here, too, identity is central. The sense of shared community requires that the participants be sympathetic to the ideas around which the group is based; even if they disagree, there needs to be some fundamental common ground. Trust in the shared motivations and beliefs of the other participants - in other words, their social identity - is essential to the sense of community [Beniger 87].

Identity also plays a key role in motivating people to actively participate in newsgroup discussions. It is easy to imagine why people may seek information on the net: they have a problem and would like a solution. What prompts someone to answer? Why take the effort to help an unknown and distant person? Altruism is often cited: people feel a desire or obligation to help individuals and to contribute to the group [Constant et al. 95]. Yet selfless goodwill alone does not sustain the thousands of discussions: building reputation and establishing one's online identity provides a great deal of motivation. There are people who expend enormous amounts of energy on a newsgroup: answering questions, quelling arguments, maintaining FAQs[1]. Their names - and reputations - are well-known to the readers of the group: other writers may defer to their judgement, or recommend

that their ideas be sought in an argument. In most newsgroups, reputation is enhanced by posting intelligent and interesting comments, while in some others it is enhanced by posting rude flames or snide and cutting observations. Though the rules of conduct are different, the ultimate effect is the same: reputation is enhanced by contributing remarks of the type admired by the group. To the writer seeking to be better known, a clearly recognizable display of identity is especially important. No matter how brilliant the posting, there is no gain in reputation if the readers are oblivious to whom the author is.

Models of honesty and deception

The approach of this paper is ethnographic - an interpretation of closely examined social discourse [Geertz 73]. As a framework for the examination I will look at the virtual community as a communication system, its inhabitants as signallers and receivers.

Examples of identity deception abound in the animal world[2]. The deception is quite harmful to those deceived, whose costs range from a lost meal to loss of life. However, it is beneficial to the deceivers, who gain food, free child care, or their own safety. What maintains the balance between honest and deceptive signalling and why, since it can be so beneficial to the deceiver, isn't deception more common? Why don't more harmless butterflies mimic the bad-tasting monarch? And why don't weak, undesirable mates just pretend to be strong, desirable ones?

There is not a simple answer to this question; there is not even agreement among biologists as to how common, or effective, is deception. If a signal becomes very unreliable due to excessive cheating it ceases to convey information - it stops being a signal. Yet there are stable systems of deception, where the percentage of deceivers does not overwhelm the population, and the signal remains information-bearing, however imperfectly. And there are signals that are inherently reliable: signals that are difficult, or impossible, to cheat.

Biologists and game theorists have developed an analytical framework for modeling the interplay between honesty and deception in a communication system. Of especial interest to us is the work done by Amotz Zahavi and others in examining what makes a signal reliable. Zahavi proposed the ``Handicap Principle'', which states that

> ...[F]or every message there is an optimal signal, which best amplifies the asymmetry between an honest signaller and a cheater. For example, wasting money is a reliable signal for wealth because a cheater, a poor individual claiming to be rich, does not have money to throw away; the message of strength may be displayed reliable by bearing heavy loads; and confidence may be displayed by providing an advantage to a rival. [Zahavi 93b]

Signals that follow the handicap principle are called **assessment signals**. They are costly and the cost is directly related to the trait being advertised. Big horns on a stag are an assessment signal for strength, for the animal must be quite strong and healthy to support these massive growths. The horns are a signal: potential rivals or mates need not directly test the stag's strength; they can simply look at the size of the horns. The thick neck of a brawny bouncer in a bar sends a similar signal in the human world and few patrons demand a personal exhibition of strength. Assessment signals are reliable, since sending an assessment signal requires that the sender possess the relevant trait.

Signals that do not follow the handicap principle are called **conventional signals**. Here, the signal is correlated with a trait is by custom or convention: the sender need not possess the trait in order to make the signal. Wearing a T-shirt that says ``Gold's Gym Powerlifter'' or signing ``Mr. Deadlift'' in your letters to a weight-lifting newsgroup is a signal of strength, but not a reliable one. Anyone can wear the shirt or type the signature: no feats of strength are involved in the signal's production. Conventional signals are open to deception. If being thought of as strong is highly desirable, it seems reasonable that many people, weak or strong, would choose to wear ``Powerlifter'' T-shirts.

Yet, if T-shirt wearing by the weak becomes prevalent, the signal loses its meaning as an indicator of strength. Conventional signals can thus be unstable: due to excessive deception, a once meaningful signal can become noise.

Since assessment signals are so reliable, and conventional ones not, why use the latter at all? One reason is that conventional signals are often less costly, for both the signaller and the receiver [Dawkins & Guilford 91]. If the costs associated with deception are relatively low, then a conventional signal may be more suitable than a more reliable, but costly, assessment signal. Think of a job applicant. The text in a resume is a conventional signal, for one can write down an impressive job history without having actually experienced it. Statements made during an extensive interview are more like assessment signals, for one must have actually acquired the knowledge in order to display it. It is much quicker and easier (that is, less costly) for the employer to just look at the resume, but the chances of being deceived are much higher. If the costs of deception are high - say, the job is a responsible one and an inexperienced employee may cause a great deal of harm - then it will be worth making the effort to make the costlier evaluation.

The spread of deception can be limited. In particular, imposing a cost to being caught deceiving - that is, punishing deception - is a deterrent [Hauser 92]. Going back to our resume example, there is usually little penalty for being caught padding one's employment history when seeking, say, a waitressing job, whereas the punishment for being caught amplifying one's medical qualifications may be quite severe. By imposing high costs on deception a social system can make conventional signals more reliable.

There are costs to imposing the punishment. For a deceptive signaller to be caught, someone must make the effort to assess the honesty of the signal (termed the *cost of probing* among biologists). In addition to the time and energy thus expended, probing may itself have high costs if the probee turns out to be honest [Dawkins & Guilford 91]. For instance, imagine probing the strength of our ``Powerlifter'' T-shirt wearers by challenging them to a fight. If the wearer is ``deceptive'', he or she will lose the challenge - and quite likely be dissuaded from misleading T-shirt practices. If, however, the wearer is ``honest'', it is the prober who loses.

Applying the model is an interpretive process. Even in the relatively simple world of biological signalling, there is often disagreement about whether a particular signal is inherently or conventionally tied to a trait (see Dawkins & Guilford 91 and Zahavi 93b for opposing views). Interpreting the social world of cyberspace is far a more subjective process. The purpose of the model is to help articulate the arguments: it is a framework on which to begin sorting out the intricate and often murky discussion about identity in cyberspace.

The anatomy of a Usenet letter

Usenet news is accessible by millions of people all over the world [3]. Subscribers range from the highly technical to the computer illiterate, from young children playing with their parents' account to homebound elderly people using the Net for social contact, from young American urban professionals to radical Afghani Muslims. Some are posting from work accounts, knowing their boss monitors on-line exchanges; others are posting from a recreational service, entertaining themselves by playing an imaginary character.

There are several hundred discussion groups, covering topics ranging from computer networking protocols to gun-control and vegetarian cooking. Some newsgroups encourage anonymous postings; in others such postings are coldly ignored. Some newsgroups are close-knit communities, in which people refer to each other by name and ask after each other's friends and family members. Others are primarily places to exchange information, repositories of knowledge where one can submit a

question and receive a (possibly correct) answer. Some groups provide a warm, trusting and supportive atmosphere, while others promote a raw and angry free-for-all.

This range of styles, topics, and participants makes Usenet an especially interesting focus for this study. Although the groups share the a common technology and interface, the social mores - writing style, personal interactions, and clues about identity - vary greatly from forum to forum.

In the mediated world of the Usenet newsgroups, the letter is not only the basic form of communication, but one's primary means for self-presentation. In the following section, we will take a close look at the anatomy of a Usenet letter and at how identity is established or concealed within it.

Example 1: *A Usenet Letter*

header	Owen@netcom.com (Owen Koslov) Re: Brine Shrimp Mon, 24 Jan 1994 09:38:23 GMT Newsgroups: sci.aquaria,rec.aquaria In article fdorrance@aol.com (F. Dorrance) writes:
Message body (extract from original question)	>> I tried to hatch some brine shrimp for my fish. I could >> only get the shrimp to live for 2 days. Could someone >> tell me what to feed them and give me details on >> hatching them.
Message body (response)	You are not supposed to keep them alive for longer than a day or so. They should be fed to the fish as soon as they hatch. Otherwise, you need the type of set up you'd expect in a regular saltwater tank: low bio-load, plenty of water circulation, and adequate filtration. You can feed the shrimp OSI's APR or other commercial invert foods, or use green water. In all cases, unless you are doing it on a large scale, buying live brine shrimp at a shop is simpler, faster and easier.
Signature	-- Owen Koslov at home (owen@netcom.com) or work (okoslov@veritas.com)

Example 1 is a typical Usenet letter. In this example, the writer is offering advice in response to a request for information. What does the reader of this posting know about Mr. Koslov [4]? Clues can be found in each part of the letter: the header provides the writer's name and email address, the body of the letter reveals voice and something of the history of the exchange, the signature shows the writer as he or she chooses to be identified.

The account name - basic ID

The most straightforward form of identification is the writer's account name (i.e. email address). This information is automatically included in the header by the posting software. It appears in the article lists Usenet readers skim to find postings of interest and it is the data used in killfiles to identify writers one finds onerous [5]. The automatically inserted account name may be the only overt identifier in the posting; while people do not always sign their letters, all postings must have the senders account name in the header.

A close look at the account name, a seemingly simple identification signal, proves to be quite interesting for it touches on issues ranging from the reputations of various virtual neighborhoods to techniques for detecting identity deception. There may be a clear and straightforward mapping from an account name to a real-world individual - or it may be deliberately opaque. The domain (account names are in the form *name@domain*, where domain is the organization that provides the account) yields contextual clues about the writer - and about the reliability of the header information.

While the name of the individual writer may be unfamiliar, often the name of the domain is not. Like notes written on letterhead, a posting submitted from a well-known site shares in its reputation: a posting about oceanography has added authority if it came from *whoi.edu* (Woods Hole Oceanographic Institute) and a question about security breaches may seem more intriguing if it came from .mil (the U.S. military).

The domain is a virtual neighborhood and, as with real world neighborhoods, some names bring to mind a wealth of demographic data. The domain may correspond to a real world place, indicating that the writer is in, say, Thailand or Israel or working at Raytheon or Greenpeace. Even some of the commercial services have distinctive reputations: San Francisco's The Well is tie-dyed and politically active while New York City's Echo is black-clad and arts-oriented. There are ``poor'' neighborhoods, addresses that reveal a limited budget:

>> jake@cleveland.freenet.edu
>> Hacker wanted to disassemble commercial program and rewrite to our specs.
>> This is not a B.S. post, we willpay BIG $$$ to have this service performed.
>> Email for details.

Nobody with ``BIG $$$'' to spend is going to be writing that message from Cleveland Freenet. :-)

And, while there are not yet any recognized ``wealthy'' virtual neighborhoods, it is probably only a matter of time until exclusive on-line addresses become symbols of status.

To understand the significance of the domain it is useful to distinguish between institutional and commercial accounts. Institutional accounts are online addresses from universities, research labs and corporations; they are given to people because of an association with the institute. Commercial accounts are available for a fee from various service providers. Unlike the institutional accounts, these commercial accounts do not imply any affiliation; they simply mean that the user has signed up for the service.

In the early days of the net, all accounts were institutional. Most sites were big universities and laboratories and the users were academics and researchers at these institutes. Today, the situation is more complex and not all postings are from recognizable institutes. Some are from small businesses unknown to the reader; others are accounts from commercial service providers -somewhat like electronic post office boxes. Furthermore, as net access becomes widespread, a posting from a research lab domain no longer necessarily means a researcher sent it. Support staff as well as scientists have computers: the posting from Woods Hole may be from a prominent oceanographer - or a temporary receptionist. There is a great deal of contextual and other information to be found in the domain name, but it needs to be evaluated within the culture of the net and of the organizations that provide access.

The opening up of the online world to anyone with a computer and modem has met with quite a bit of resistance from the original residents. Most maligned are newcomers who have accounts with the consumer-oriented commercial services such as Prodigy and America Online (AOL). Postings with aol.com addresses are sometimes greeted with derision; newsgroups such as alt.aol-sucks exist solely to spurn America Online subscribers. This resistance is partly a reaction to the loss of exclusivity - access to online communication no longer means one is at the forefront of technology - but there are also substantive differences between the postings of the old guard and the newcomers.

I have to admit, i do have an account on prodigy, becasue my mom has had it for a few years, and it was free for me, i never use it though, it's embarassing.. I think i'd be the same way if i had AOL, i would be embarassed to post or reply or just be seen anywhere with a loser@aol.com address. I guess it just seems that they are so stupid..

- alt.2600

Some of the differences are stylistic. The consumer-oriented services offer their own communication forums in which the conversational conventions are quite different than within Usenet. For instance, in an AOL chat room, it is fine to simply respond ``yes'' to a statement. On Usenet, where each statement is a stand-alone posting, it is considered poor manners to post a response with no added content and with no indication of the original statement. Newcomers to Usenet from AOL, accustomed to chat room style interactions, frequently post one word rejoinders, infuriating other Usenet users and adding to the image of AOL-based participants as thoughtless and ignorant of local customs.

There are also differences in accountability. The holder of an institutional account - a student or employee - has reasons, such as a job or degree, for remaining in good standing with the account provider. A user who engages in malicious or illegal activity online stands to lose more than just the account: a number of students have been disciplined - and some expelled - for violations of institutional policy. The relationship of a subscriber to a commercial service is much less consequential. While most services have policies about what constitutes acceptable behavior, the repercussions for infractions are limited to termination of service - an inconvenience, certainly, but not the equivalent of demotion or firing. In other words, higher punishment costs can be imposed on the institutional account.

Institutional accounts are also less private than commercial accounts. A work- or school-based account name is known within the organization and there are many people who can make a direct connection between the name on a posting and the real world person. For the writer using an institutional account, the online world is a public forum in which he or she can be seen by numerous colleagues and acquaintances. With a commercial account, it is up to the user to decide who should know the link between physical self and virtual appellation. Some services allow subscribers only a single account name and the user thus has some concern about the reputation that attaches to that name. Some consumer-oriented commercial services make it easy for subscribers to create multiple, fictitious names and to keep their real names from appearing on their postings. The anonymity of these accounts makes them popular for disruptive and harassing posting.

Truly anonymous postings can be sent using anonymous remailers. These are forwarding services which will strip all identifying information from a letter and then forward it, anonymously or under a pseudonym, to the intended recipient or newsgroup. The pseudonymous address added by many, though not all, remailers clearly indicates that the posting is anonymous: *an12321@anon.penet.fi* and *anon-remailer@utopia.hacktic.nl* are typical. While remailers can be used maliciously, their primary use is to provide privacy. Anonymous posts are common in groups where the participants reveal highly personal information and many of the support groups (e.g. *alt.support.depression*) periodically provide instructions on how to use an anonymous remailer. Use of a remailer can also be a political statement, an affirmation that one supports citizen's right to privacy (which includes anonymity, access to strong encryption tools, etc.) and opposes government and corporate surveillance.

Many Usenet participants frown upon anonymous postings:

> I don't like dealing with anyone that uses an anonymous remailer! I will, however, assume you are doing this for 'legitimate' reasons and try to render assistance.
>
> *- comp.unix.security*

> The anonymous address and excessive crossposting were a bit much don't you think? I'll humour you this time...
>
> *- alt.2600*

These writers felt compelled to point out their disapproval of anonymous posting before answering the question. And the stigma of writing anonymously was clearly felt by *an451494@anon.penet.fi*,

the writer of an innocuous request for reviews of a human interface, who signed his posting with his real name and the note ``sorry; my employer doesn't like seeing me posting to news-groups''.

The account name is thus an important, but limited, form of online identification. It is important because it is ubiquitous: all postings must have the account name in the header. It is a key marker of individual identity: although there is not always a one-to-one mapping between an account name and a real world person (accounts may be shared, some people have several accounts), the account name is generally perceived to refer to a single person (or persona). And it may provide some contextual information about the writer, information that, while quite sketchy, may be the only such cues in the posting.

Identity in voice and language

The contents of the posting can reveal a great deal more about the writer. It may include overtly identity-related data: name, age, etc. More importantly, it provides a chance to get a sense of the writer's ``voice'' and to see how he or she interacts with others in the on-line social environment.

Erving Goffman, in his classic work The Presentation of Self in Everyday Life distinguished between the ``expressions given'' and the ``expressions given off''. The former are the deliberately stated messages indicating how the one wishes to be perceived; the latter are the much more subtle - and sometimes unintentional - messages communicated via action and nuance [Goffman 59]. Both forms of expression are subject to deliberate manipulation, but the ``expression given off'' may be much harder to control. One can write ``I am female'', but sustaining a voice and reactions that are convincingly a woman's may prove to be quite difficult for a man.

Is the ``expression given off'' an assessment or conventional signal? Even in the real world, with its far richer array of social nuance, imposters exist. Drag queens, confidence men, undercover spies - all are adept at mimicking subtle social codes. Yet, although nuance is not an infallible indication of a social role, the experience of living the part greatly influences one's ability to play it: years of socialization make most women adept at playing the role of woman in their culture and one can argue that the cost in time that it takes to fully attain the role makes it an assessment signal [6]. Is the same true in the virtual world - is there sufficient complexity to the nuances in a Usenet exchange for one's experience in living a role to be revealed in the ``expression given off''?

Looking again at Example 1, from the header and signature we know the writer's name (Owen Koslov), we know that he writes both from home (as *owen@netcom.com*) and from work (*okoslov@veritas.com*). Neither location tells us a great deal about the author - netcom is a commercial service provider and veritas.com is not a well-known company. The writer's history on the net reveals much more. A look at recent articles shows that he is a fairly frequent writer, not only on sci.aquaria, but also on the closely related groups rec.aquaria and alt.aquaria. Indoor aquaria seem to be his passion. He provides a fellow killifish fancier with the address of a mailing-list devoted solely to killifish. He writes several letters a day on aquaria related topics. We learn that he has perhaps too many fish: ``I wish I had your discipline in keeping the number of species down. I have 9 species of lampeyes alone...'' His letters are usually answers to questions posed by others, his voice is usually authoritative, pedantic, occasionally dryly humorous. Here are selections (ellipses mine) from his response to someone who said to avoid charcoal in tanks with plants.

> While activated carbon does adsorb more than just organic carbons, a categorical statement like that is inaccurate. Carbon may remove some trace minerals, but I challenge you to substantiate the statement that it is a ``bad thing'' for live plants. ... Further more, Dick Boyd's Chemi-Pure uses activated carbon as one of its ingredients and I am yet to hear one credible report of it negatively affecting live plants. The late Dr. Bridge had used a mix of activated carbon and peatmoss as a filtration medium for his planted show tank and reported excellent results.

Over time, the frequent contributors to a newsgroup creates a strong impression. The reader of *rec.aquaria* is likely to be familiar with these postings and has come to some conclusion about both Mr. Koslov's reliability and his personality. Although this writer says little about himself, there is a great deal of expression given off.

Writing style can identify the author of an posting. A known and notorious net personality hoping to appear online under a fresh name may have an easier time disguising his or her header ID than the identity revealed in the text. The introduction to the cypherpunks newsgroup includes this warning:

> The cypherpunks list has its very own net.loon, a fellow named L. Detweiler. The history is too long for here, but he thinks that cypherpunks are evil incarnate. If you see a densely worded rant featuring characteristic words such as ``medusa'', ``pseudospoofing'', ``treachery'', ``poison'', or ``black lies'', it's probably him, no matter what the From: line says.

> *- Cypherpunks mailing list*

In this case, where the usual assessment signal - the name in the header - is believed to be false, language is used as a more reliable signal of individual identity.

Language is also an important indication of group identity. ``[R]egarding group membership, language is a key factor - an identification badge - for both self and outside perception.'' [Saville-Troike 82] Language patterns evolve within the newsgroups as the participants develop idiosyncratic styles of interaction - especially phrases and abbreviations. Some are common to all groups: BTW, IMHO, YMMV (By The Way, In My Humble Opinion, Your Mileage May Vary). Others are of limited extent: MOB, ONNA (Mother Of the Bride, Oh No Not Again - used in *misc.kids.pregnancy* by women who were trying to get pregnant to report the monthly disappointment). New words are coined and ordinary words gain new meaning: flame, spam, troll, newbie. Using these phrases expresses ones identification with the online community - it is akin to moving to a new region and picking up the local accent.

Participants interpret these language cues according to their own position within the social group. A newsgroup can be home to two or more factional groups, each of which tries to establish its style and views as the rightful culture of the group. Alt.2600 is a newsgroup devoted to computer hacking and related topics. Being a hacker - or appearing to be one - has recently gained mass popularity, and would-be hackers and trend-conscious teens have adopted a style of writing that features alternative spellings (such as ``kewl'' for ``cool'') and random capitalization. Older or more experienced hackers felt compelled to separate themselves from the crowd. In a thread called ``Attn LaMerZ and Wana-B's'', one wrote:

> ATTENTION you are not a hacker if you have seen the movie HACKERS. ATTENTION you are not a hacker if you post here looking for AOLHELL, free AOL, Unix passwords, crackerjack, or virus creation lab. ATTENTION you are not a hacker if you HaVE A PrOOBLEM WiTh YOuR CaPS LOckS KeY.

> *-alt.2600*

Here, the language markers one group developed to distinguish themselves are a sign of scorn for the other.

The signature

Language markers such as the above are an important element in signatures, which are the on-line world's most deliberate identity signals. The signature is added at the discretion of the user, though once designed, it is usually appended automatically to postings. It may be an electronic business card, an elaborate work of self-expression, a cryptic remark, or simply a name. Not everyone uses one, and they are far more prevalent in certain forums than others. Although the signature itself is an

easy to copy conventional signal, it is often used as a means to link to more robust and reliable indicators of identity and to show writer's the affiliation with a subgroup.

Signatures can be used to anchor the virtual persona to the real world person. The net is a great leveler: no one knows if you are male or female, boss or underling, gray-haired or adolescent; ``on the Internet, nobody knows you're a dog". This is not to your advantage if in the real world you hold some authority: no one can see that you are a respected professional at work in your office, not a teenager logging in from a bedroom. One use of the signature is to present real world credentials: your full name, title, department, office phone number; enough information so that someone could, if they were curious, check to see that you were really who you claimed to be. Such business-card signatures are common in the technical newsgroups. Advice from someone who's job title is ``Unix System Specialist" or ``Director, Software Development" has added weight, particularly if it is in a known and respected company (for the important sounding ``Director of Software Development" at unknown ``ABC Software Co." may be also be the founder and sole employee). These signatures imply that the writer is posting in his or her official, employed capacity - willing to publicly stand behind the statements.

One newsgroup that contains many business-card signatures is *comp.security.unix*. The discussion here is about how to make unix systems secure - and about known system flaws. While many of the participants are system administrators of major institutions, others are just learning how to set up a system in a fledgling company and some, of course, are hoping to learn how to break into systems. Someone posting a question may wish to include credentials to assure potential responders that the question is legitimate, not a disguised dig for information from a would-be hacker. A posting suggesting that administrators improve their sites by changing this or that line of code in the system software could be a furtive attempt get novice administrators to introduce security holes; a signature verifying the legitimacy of the writer alleviates this suspicion. Identity deception is a big concern of the participants in this group, and they are very aware of signatures and their implications.

An important new use of the signature is to refer to the writer's home page on the World Wide Web. Like the business card signature, the Web address may contain credentials - and much more. A home page may provide a detailed portrayal of its subject: people include everything from resumes and papers to photographs and lists of favorite foods. A person's presence on the Web has depth and nuance not found in the ephemeral Usenet environment and a writer's self-presentation on the Web can provide a very enlightening context for understanding his or her postings.

Signatures often include a disclaimer, saying something to the effect that ``These are my opinions and not those of my employers". For many people, participating in Usenet newsgroups occupies a sometimes awkward position between work and private life. The newsgroups may be an important resource for one's work; they may also be a purely recreational past-time. Whether a posting is about signal processing or Argentinian culture, if sent from one's work account it will show up under the company's electronic letter-head. The disclaimer proclaims that the writer is appearing as himself, not as an official company spokesman.

Signatures are also used to establish one's ties to online groups. Many signatures contain computer jokes and phrases, showing that the writer is a programmer - a member of the old guard of the net.

Example 2: *Programmer signatures*

    a. #include <stddisclmer.h>

    b. Write failed on /dev/brain: file system full

    c. Doom: 5% Health, 0% Armor, 59 cent Tacos, Lets Go!

    d. Dave Mescher dmescher@csugrad.cs.vt.edu
       GCS d H>+ s+:- g+ p3 au a-- w+ v,--->! C++++,++ UU++++,

A$ P-- L-3- E--- N++ K- W--- M V-- po Y+ t--- 5 jx R G+
!tv b++ D- B--- e+,* u+ h- f+ r(+,++)@ !n,--- y?

In a), the phrase #include <stddisclmer.h> uses a C language construct to make a reference to the disclaimer signature. It has both the effect of being a disclaimer (the writer is not speaking officially) and of proclaiming the writer's affinity to the C programming world. b) uses the format of a common Unix error message. For the Unix-literate reader, the phrase is familiar and the joke obvious; for others it is simply obscure. Similarly, c) plays off the scoring style of the popular computer game Doom. Signatures such as these are often individual creations, meant to be used only by their author. Since one needs to be familiar with a subculture in order to make a joke in its vocabulary, these signatures, when original, show their author's familiarity with the programming world. Furthermore, the world in which they are used is small. A writer who simply copies the clever phrase of another is likely to quickly come to the attention of the potentially irate creator. The possible cost of copying - public humiliation through accusations of plagiarism - is quite high. (Smitten by a phrase too witty not to use, some writers have taken to using other's signatures - with attribution.).

These signatures are what Fiske calls ``producerly'' writings, easily accessible yet playing with complex mixes of vocabularies and codes [Fiske 89]. Such puns, he says, entertain both through the process of discovering the layers of meaning and in their juxtaposition of social contexts. ``Write failed on /dev/brain: file system full'' is not only a play on Unix error messages, but can also be read as a comment on information overload - or as a subtle jab at those who post, but seem not to absorb anyone else's comments. ``#include <stddisclmr.h>'' in addition to mixing the culture of Usenet with the code of C programmers, also refers to the numbing ubiquity of disclaimers.

The signature in d), which includes the writer's ``Geek Code'' is a bit different. Proclaiming one' s ``geek identity'' - both one's identification with geeks as a group and one's particular and individual type of geekiness - is the purpose of this code:

> How to tell the world you are a geek, you ask? Use the universal Geek code. By joining the geek organization, you have license to use this special code that will allow you to let other un-closeted geeks know who you are in a simple, codified statement.

-Robert Hayden, The Code of the Geeks, 1995.

The Code consists of a series of descriptive categories and modifiers. The first category is G, Geek type. GCS stands for Computer Science Geek (GSS would be Social Science Geek). The second category, d, is for dress style. In the example above d is without modifiers, meaning: ``I dress a lot like those found in catalog ads. Bland, boring, without life or meaning''. There are many possibilities, ranging from ``d++: I tend to wear conservative dress such as a business suit'' to ``!d: No clothing. Quite a fashion statement, don't you think?''. The Geek Code can become quite complex. For instance, the modifier ``>'' means moving towards. Thus, the symbol H>+ above is interpreted as someone whose hair (H) is striving to achieve (>) shoulder length (+). The Code is full of inside jokes, e.g. !H, the code for baldness, refers to the computer language convention of ! meaning negation. It is also full of cultural references: to operating systems, to Internet personalities, to TV shows and to various games. One has to be quite dedicated to decipher a Geek Code, but its primary message - identification with the online ``geek'' world - is easily perceived by anyone who knows what a geek code looks like.

The Geek Code has inspired a number of other identity codes. The Goth Code has categories for dress style, body piercings, musical taste; the Magic Code provides the means to express one's opinions of the Kabbala, Aleister Crowley, and one's own supernatural powers; the Cat Code has categories ranging from breed to purr volume. The Codes seem to have originated in the gay and lesbian on-line community, inspired both by the handkerchief codes of the gay bar scene and (according to the introduction to the progenitor of all the codes, the Bear Code) by the astronomical classification schemes for stars and galaxies. Until recently, seeing these codes in a signature would

be a puzzle to all but the initiated. Today, the Web has made finding esoteric information - such as the decoding scheme for the Muffdiva Code - simply a matter of a quick net search. The codes still function as subcultural membership markers, though their meaning is now open. They can now function as a tourist's introduction to the subculture, enumerating the features of greatest interest to the group.

Some signature styles are unique to particular newsgroups. Often they refer to the writer's role within that group. For example, *soc.couples.wedding* is a newsgroup devoted to planning weddings from the fine points of invitation writing protocols to advice about how to deal with hostile in-laws. The participants include people who are engaged to be married, people who would like to be engaged to married, and people who like giving advice. The brides-to-be are the central group: they ask questions, they share their experiences, they write to complain about their caterers, bridesmaids and future mother-in-laws. And they have developed their own signature pattern:

> Joan (and Mike, May 27, 1995)
> Amy (& Chris Sept. 7, 1996)

A similar signature pattern is found on *misc.kids.pregnancy*, where the expectant mothers sign with the baby's due date. These signatures show the special status of the writer: as bride- or mother-to-be, her real world situation is the focus of the group's interest. The signatures also highlight the temporal nature of this identity. The readers know the stages of wedding preparation and pregnancy. Responses to letters often include references to the signature (``June 10th - I'm getting married the next day! Are you nervous yet?'') even if the body of the letter was unrelated to the writer's wedding.

An especially well-defined community has emerged in the group *rec.motorcycles*, where an online club, called the Denizens of Doom (DoD), has formed. The DoD began as a satire of the newsgroup and of real motorcycle clubs, but is now a real club, with memberships lists and real-world badges. Members get DoD numbers, which they use in their signatures. One must apply for membership, and, while the procedure is not terribly secret, it does take a bit of knowing who's who in the group to apply: a DoD number in a signature means that the writer is not a newcomer to the group. Here is a signature from rec.motorcycles:

> whiteb1@aol.com (Ben White)
> AMA # 580866 COG # 1844
> DoD # 1747 Better watch out, He turned me loose!
> '95 VFR 750 5 bucks more, I coulda got a red one
> '85 Shadow
> No more Connie

In addition to the DoD number, it features the writer's membership in the American Motorcycle Association plus the motorcycles he owns (or used to own); this signature is a virtual world substitute for the colors and badges of real-world biking.

Finally, signatures make it easier to quickly identify the writer. In the uniform environment of ASCII text, there is little to visually distinguish one letter from another and it is easy confuse two writers with similar names, or to simply not notice the attribution at all. Signatures are easily recognized, identifying the writer at a glance.

Individual recognition is important in many newsgroups. Participants in arguments often call each other by name - both heated flames and supportive letters are often written as person-to-person missives. On-line status is recognized and there is deferral to respected members. This writer, himself an aerobics instructor, described a modification he had made to a move, and then asked:

> Bill Whedon and Larry DeLuca, Are you There? You guys have seemed to be the most vocal AND
> concientious... By turning lunges into squats, have Lori and I traded one problem for another?

- misc.fitness.aerobics

High status participants get special treatment. A bridal consultant who contributes frequently to *soc.couples.wedding* asked the newsgroup to help her plan a vacation. Such a request is quite outside the group's domain, and would normally result in a sharp requests to keep the postings on topic; instead, several people enthusiastically wrote in with advice and suggestions. The signature is an important technique for insuring that one's postings are accredited to one's name.

From the header to the signature, identity cues are scattered throughout the Usenet letter, from declarations of one's name, age, sexual orientation, to the subtler ``expressions given off'' through voice and vocabulary. The virtual world's subcultures have developed their own patois, with codes and linguistic patterns that identify affiliated participants. And people have found ways to control the degree of personal identity they wish to expose online, from authentication through business card signatures to the private cloak of anonymous remailers.

Deceptions and manipulations

Yet these identity cues are not always reliable. The account name in the header can be faked, identity claims can be false, social cues can be deliberately misleading.

Many varieties of identity deception can be found within the Usenet newsgroup. Some are quite harmful to individuals or to the community; others are innocuous, benefitting the performer without injuring the group. Some are clearly deceptions, meant to provide a false impression; others are more subtle identity manipulations, similar to the adjustments in self-presentation we make in many real world situations.

Trolls

In the spring of 1995 a new user appeared in the wedding newsgroups. She signed her letters Cheryl, the name on her account was *ultimatego@aol.com* and her letters espoused a rigid interpretation of formal etiquette. The discussion in these groups is often about how to have a wedding on a limited budget. When the women would talk about using balloons for decorations, Ultimatego would post that balloons were vulgar; when the discussion turned to do-it-yourself laser-printing she would interject that only engraving is acceptable to people with taste. Some readers were intimidated by her intimations of upper-crust social knowledge; others were infuriated by her condescending remarks. When she wrote that people who could not get married in full formal splendor should not have a wedding at all but should simply go to city hall, an intense and angry exchange ensued. At this point, someone said that Ultimatego was probably a troll.

> Are you familiar with fishing? Trolling is where you set your fishing lines in the water and then slowly go back and forth dragging the bait and hoping for a bite. Trolling on the Net is the same concept - someone baits a post and then waits for the bite on the line and then enjoys the ensuing fight.

Trolling is a game about identity deception, albeit one that is played without the consent of most of the players. The troll attempts to pass as a legitimate participant, sharing the group's common interests and concerns; the newsgroup members, if they are cognizant of trolls and other identity deceptions, attempt to both distinguish real from trolling postings and, upon judging a poster to be a troll, make the offending poster leave the group. Their success at the former depends on how well they - and the troll - understand identity cues; their success at the latter depends on whether the troll's enjoyment is sufficiently diminished or outweighed by the costs imposed by the group.

Trolls can be costly in several ways. A troll can disrupt the discussion on a newsgroup, disseminate bad advice, and damage the feeling of trust in the newsgroup community. Furthermore, in a group

that has become sensitized to trolling - where the rate of deception is high - many honestly naive questions may be quickly rejected as trollings. This can be quite off-putting to the new user who upon venturing a first posting is immediately bombarded with angry accusations. Even if the accusation is unfounded, being branded a troll is quite damaging to one's online reputation.

Rec.motorcycles is a free-wheeling group where tough-guy banter mixes with advice about riding techniques and equipment. Being able to ride on a challenging bike in difficult conditions is respected - but attempting feats beyond one's capabilities is greatly disapproved of. Beginners who want to start out on a powerful bike are likely to be severely lectured. Provoking this response is the goal of flame-seeking trolls.

> Subject: New Rider; what bike? Is ZX11 good to start with?
> From: crystllthr@aol.com (CrystlLthr)
> Organization: America Online, Inc. (1-800-827-6364)
>
> Hi. I am a college junior and am interested in buying my first motorcycle. I've seen the Kawasaki ZX11 and think it looks pretty hot. Would this be a good bike to buy. Money is no problem. My dad will buy me anything I want. Also, I've heard that you should get a turbo kit from Mr. Turbo in Houston, because the bike needs more power. Any other modifications suggested? Also, where should I go to learn to ride?
>
> Derick Nichols
> Tulane University

A few took this posting seriously. Some responded angrily:

> Cool Derick. Great bike! The turbo kit will make it awesome. Don't worry about learning to ride, just go pick it up and ride it home. Good idea though to have your dad buy a coffin and funeral plot at the same time. I think they offer a deal on those when you get a zx-11 as your first bike. Oh yeah. Don't worry about hitting any manhole covers on your way home - if you warp the wheel you can always sue the city.

Others tried to be helpful:

> Well, I agree with Sherry [quoted above].though i problby wouldnt have put in quite like that. Derick, if you have never ridden before maybe you can start with something a bit smaller. You dont hae to buy your first bike nwe.,that way, when you otgrow it, you wont have put a lot of money into it (if you bought it cash), or if you financed, you can get out ot it easier.

Most readers, however, decided it was a troll:

> This has got to be bait, right?????
>
> Sounds like extreme flamebait to me..
> Yes, get the ZX, (a used '91.) I'll buy the resulting wreck for parts for $100.00 :-)

Several pointed out the discrepancy between the signature and the domain:

> >> This has got to be bait, right?????
> Since the "college junior" is not coming from an .edu address, then I would guess yes it does. :-)
>
> College junior @ aol.com? Bad bait, too obvious.

Still, as with many issues of online identity, the question of Derick's intentions remained unresolved:

> >> >>I worry about people like this on motorcycles.
> >>I worry about people like this behind on computers. No matter how lame the bait someone will bite.
> What scares me even more is teh possibility that it isn't bait....

In a group such as *rec.motorcycles* an occasional troll is not too harmful. The troll's game of testing the participants' astuteness is not too far in spirit from the newsgroup's normal banter and remarks such as ``Bad bait, too obvious'' imply the testing goes both way. A better troll would be admired for cleverness; the offense here was not trolling per se, but doing it so poorly.

In other groups the presence of a troll can inflict quite a bit of harm by undermining the trust of the community. The wedding newsgroups that Ultimatego frequented consist of women (and a few men) from very disparate backgrounds discussing the planning of a highly emotional event fraught with concerns about family, tradition, money, and status. The culture that has evolved frowns upon any authoritative statement of the ``right'' way to do things and writers frame their advice with phrases such as ``in my opinion'' or ``it is often done this way''. Ultimatego's early posts were not overtly offensive, but their formal and imperative voice was at odds with the conversational tone of the other participants.

One woman wrote:

> Hi Everyone, Some of my coworkers and I were wondering if it is still considered a faux pas to wear white to someone else's wedding. One of the girls just got married and said she noticed that about 20 women (400 guests) wore white outfits. It didn't bug her, but another guest commented about it later. thanks, Jaime (and Jet) 03-09-96

Ultimatego responded:

> Dear Jaime,
>
> It is consider improper to wear white at a whedding, since it appears to compete with the bride. The guests were improperly dressed. It is taboo to wear black...even as a fashion statement since it is associated traditionally with mourning.
>
> Kindly,
> Cheryl.

Had Ultimatego maintained this persona, it is quite possible that she would have been accepted as the wedding groups' duenna, uninvited but not entirely unwelcome. However, Ultimatego's facade kept slipping. She went from chilly to rude, her proper grammar sliding into vicious name calling. Participants who could otherwise count on a generally supportive reaction from the group found themselves subject to Ultimatego's attacks. To someone seeking advice about a painful issue, such as a parent who is refusing to come to the wedding, the feeling that part of the audience was motivated by hostility or perverse humor was inhibiting.

Some trolls post deliberately misleading information. In *rec.pets.cats* a writer named keffo suggested deterring cats from clawing furniture and chewing on wires by spraying hydrogen peroxide at them. Again, the readers' reactions were mixed. *Rec.pets.cats* has had a great deal of experience with Usenet pranks and many readers immediately cried ``troll!''. Others believed that she was well-meaning and simply did not know that such a technique would be extremely painful to the cat; the fact that she claimed to be a girl in 8th grade helped to explain her naivete. And some readers thought it was a reasonable suggestion, at least until a number of more knowledgeable ones explained the danger to the cat's eyes. Although past experience had taught the readers of *rec.pets.cats* that ignoring hostile posts was the best approach, this case was a bit different:

> Personally I find her type of marginal sadism towards cats as disturbing as the more overt alt.t*steless stuff, in that she could actually convince people that her suggestions are harmless. Do people who don't wear contacts know how painful it is to get the wrong solution in their eyes? I think this type of insidious troll needs stamping on as much as any other, with the proviso that in this case it is important that other people quickly point out the cruelty of her suggestions; other trolls, I think, should be ignored completely.

Responding to a troll is very tempting, especially since these posts are designed to incite. Yet this is where the troll can cause the most harm, by diverting the discussion off the newsgroup topic and into a heated argument. Instead, most groups advise ignoring such posts, both to keep the discussion topical and in the hope that, if ignored, the troll will go away. Several point newcomers to FAQs that explain how to use a killfile, which is a filter that allows one to avoid seeing any postings by a particular person or on a given topic. (Indeed, an extensive description of killfile techniques in a group's FAQ is a kind of virtual scar-tissue, an indication that they have had previous trouble with trolls or flame-wars.)

Rebuking the offending writer privately through email is also often recommended, for it does not derail the group's discussion. Such a response can be quite effective in stopping someone whose goal was not primarily to annoy others:

> I would like to apologize, to any and all of you who downloaded the junk I posted... Again I am sorry and will be more carefull in the future. Oh and I would like to thank the hundreds of you sent me E-Mail bringing it to my attention.
>
> *- comp.cad.autocad*

though it may encourage the troll whose intent was to inflame. If, however, the writer uses a false name and address, such contact is not be possible.

Responding to a troll can be costly. One may be unpleasantly insulted, as happened to this person who tried to explain to keffo (who, it later transpired, was actually a male university student) the error of her ways:

> I have tried twice to communicate with Kristen (Keffo) (by e-mail). I have found that she is nothing but a foul mouthed, uneducated, little girl who lacks respect for anything Adults;Education; Culture; Life; etc. I have received nothing but insulting profanity from this child. I explained that IF she wanted to be accepted by this group that she should issue a **Blanket Apology** to the group -- I was told what to do with my ``Blanket Apology''."

Others who have responded in person to newsgroup harassers have been mail-bombed or have had their own system administrator - or boss - contacted and told that they were making trouble on-line.

Contacting the offender's system administrator is usually done as a last resort, when it is clear that the rules of Usenet etiquette have been transgressed. Mail to keffo's postmaster complaining about the increasingly hostile postings resulted in the account being closed:

> Actually, keffo is a male university student. I've had enough of these complaints though, so don't expect to hear from him again. This is the last straw.

How seriously the system administrators or other authorities take such complaints varies greatly from provider to provider. Some may do nothing; others may be very quick to expel a user based on even a spurious complaint.

Category deception

Our perception of others is not one of wholly unique individuals, but of patterns of social categories. Our first impressions, based on brief observation, determine the basic social categories in which we place the new acquaintance, and which shape our subsequent and more detailed interpretations of their motives and behaviors [Simmel 71]. It can take significant evidence to change this initial categorization - we are more likely to reinterpret the events than to re-evaluate the basic classification [Aronson 95].

The troll is engaging in category deception. By giving the impression of being a particular type - a conservative etiquette zealot helping brides avoid errors in taste, a young girl sharing her discoveries about cat care, an earnest college junior shopping for his first motorcycle - the troll manipulates the readers' initial interpretations of his or her postings. Only when the contradictions between the troll's actions and the expectations raised by the category assessment strongly conflict does the deception begin to unravel; when, in Goffman's performance metaphor, the troll speaks out of character. Still, many readers attempt to reinterpret the actions rather than disbelieve the identification. The decisive moment in the group's realization that the postings are coming from a troll is when someone offers evidence that the real person behind the virtual identity is at odds with the one presented.

There are many other varieties of on-line category deception. Gender deception [O'Brien 97, Turkle 95] is the classic one, especially in the MUDs and in chatrooms where sex is a predominate topic of conversation (or at least, a very significant subtext to the discussion). In the Usenet newsgroups, gender deception appears to be much less common, except in forums where sex and gender are the main conversational topics. Similar category deceptions, e.g. age deception, do occur in Usenet; however, since many cases are not obvious, it is impossible to know how often or to what degree this occurs.

What does seem to be quite common here is status enhancement. Many newsgroups have some exemplary model: the consummate hacker in *alt.2600*, the cool biker in *rec.motorcycles*, the well-built body in *misc.fitness.weights*. The participant who tries to pass as an incarnation of the ideal is closely examined by the others in the group. Status in these groups is prized and, for those who claim on it is legitimate (or who have quite thoroughly deceived the others), accepting claims of dubious provenance would lessen the value and exclusiveness of their own position.

> WE all know Barry is a pathetic puppy, but his recent freaking out about my posting a picture of me has really set him off. I have many times told him to post a picture of himself so he can put up or shut up about ``how ripped an huge'' he claims to be.
>
> *- misc.fitness.weights*

The verbal claim of being muscular is a conventional signal. A change in the environment (the advent of the Web) has made it possible to send a more reliable signal of muscularity - a photograph. A prominent member of the group (the author of the above quote) put his photograph online, thus strengthening his claim to status (the photo is very impressive). Most participants applauded this effort, saying that they found it helpful to see what another participant looked like and reassuring to know that this writer's claims of expertise were indeed backed by his appearance. ``Barry'', however, greeted it with a great deal of hostility and renewed claims of his own strength; he was then challenged by many others to prove his words - to back up the conventional verbal signal with the assessment signal of a photograph.

In some groups the postings themselves are assessment signals for a salient trait. In *rec.arts.poetry*, aspiring poets submit their verses for critique by their peers; in *comp.lang.perl* programmers provide elegantly coded solutions in response to requests for help. An interesting example is *alt.hackers*. This is a moderated group, meaning that postings cannot be submitted directly to the newsgroup, but must be sent to the moderator, who (in ordinary moderated groups) filters out irrelevant or otherwise unacceptable material and posts the rest. *Alt.hackers* uses this mechanism, but without any actual moderator: to post to the group one must be able to figure out how to hack the news system. In such groups status claims and posturing are far less pervasive. This is especially noticeable in the contrast between *alt.hackers* and it's non-assessment analogue *alt.2600*. Posts to the former tend to be on topic and informative; those to the latter are often (when they are not completely off-topic) escalating boasts about petty criminal prowess.

> Why is it that everyone keeps posting that they can do amazing things,yet no proof has surfaced? Making up all this is easy because no one can rebute it, yet no one has confirmed these either. Don't post this crap

unless you have some way of confirming it.

- alt.2600

Impersonation

Not all on-line deception is involves categories. Individual identity - one's claim to be a particular individual, either in the physical or the real world - can also be challenged. A particularly costly form of identity deception is impersonation. If I can pass as you, I can wreck havoc on your reputation, either on-line or off.

Compared to the physical world, it is relatively easy to pass as someone else online since there are relatively few identity cues [7]. A surprising number of impersonated postings are made simply by signing the target's name, without copying the writing style or forging the header information. Even more surprising is how successful such crude imitations can be. Readers may pay little attention to the header information - or they may encounter the forgery in a subsequent posting, quoted without the header.

How harmful are such impersonations depends upon how defamatory the faked postings are and whether readers believe the false attribution. When impersonations are made in a newsgroup in order to discredit one of the group's participants, the target is likely to notice and post a denial:

> I am very disturbed to find that after only two weeks of Internet use, I am already being 'impersonated' by another user... Is this really so easy to do? Or did this person have to work at it? I discovered postings to a Newsgroup that appeared to come from me...but which in fact did not

- alt.alien.visitors

Since Usenet postings are not necessarily read in sequence some readers may read the forgery and miss the one that reveals the deception. This is especially likely if the faked posting set off an acrimonious flame-war: many readers will simply skip the rest of the thread and any subsequent postings made by the apparent participants.

Some of the most harmful impersonations are done without deliberate malice towards the victim, who may simply have inadvertently provided a useful identity for the impersonator to hide behind. New computer user are warned to guard their passwords carefully and to be sure to log off of public terminals. It can be very difficult to prove that one did not actually write the words that are clearly traceable to one's account.

> I would like to think everyone for bringing to our attention the outrageous message that was posted to this group several days ago... After talking with the owner of the account that generated this post. It has become clear that this is a case of a new user leaving the terminal before logging off. I ask that you please refrain from sending mail to this user (ST40L) regarding the post. He is shaken from the incident and has learned a valuable lesson the hard way.

- soc.culture.jewish

In this case, the user convinced the administrators that the posting (an anti-Semitic letter) was forged. Still, he must deal with the fact that a message that he finds abhorrent went out under his name. (This is particularly unfortunate since the debut of searchable Usenet archives. A search for letters written by this user will turn up the forgery, but not clarifications written by others, such as the one above.)

Identity concealment

Many individual identity deceptions are acts of omission, rather than commission; they involve hiding one's identity.

Sometimes identity is hidden to circumvent killfiles. Killfiles are filters that allow you to skip unwanted postings: if you put someone in your killfile, you will see no more of their postings. While killfiles may sound like the electronic version of the ostrich putting its head in the sand, they are said to be very effective in keeping a newsgroup readable. Those using a killfile no longer see the offending posts and are not tempted to respond, thus lowering the number of angry, off-topic postings. To the person who has been killfiled, Usenet becomes a corridor of frustratingly shut doors: one can shout, but cannot be heard. Some writers, determined to have their say, continuously switch the name under which their postings appear.

> I gave up trying to killfile Grubor and his myriad aliases when my filter file exceeded 10k. I am not joking about this. Admittedly only half was Johnny-boy; the other half was phone-sex spams. Still, 4-5k just on one person is a little ridiculous.

> *- news.groups*

The killfile program looks for the account name in the header, which is usually inserted automatically by the posting software. The reason someone can create ``myriad aliases" has to do with the transformation of the header from an assessment to a conventional signal.

Until recently, header information was quite reliable. Most people accessed Usenet with software that inserted the account name automatically - one had to be quite knowledgeable to change the default data. Today, many programs simply let the writer fill in the name and address to be used, making posting with a false name and site is much easier. The astute observer may detect suspicious anomalies in the routing data (the record of how the letter passed through the net) that can expose a posting from a falsified location[8]. Yet few people are likely to look that closely at a posting unless they have reason to be suspicious about its provenance.

With the header data becoming a conventional signal, such deception may be quite wide-spread. There are many benefits to using a pseudonym on-line and, unless the writer is imposing a cost on the group (i.e. being a nuisance or impersonating another participant) there is little reason to pay the costs of verifying each posting.

People have many reasons for not wanting their real names to be revealed online:

> As far as letting you know my name or giving you my fingerprints or whatever else you demand, no I don't think so. There is more going on in this net than just misc.fitness.weights. I'm involved in the net war in alt.religion.scientology. Those cultists have so far raided 4 of their net critics on bogus copyright violation charges, and in one case they placed a large amount of LSD on the toothbrush of a person who was raided, a couple of days before he was to undergo a video deposition. In my city they have been convicted of several crimes, including infiltrating the municipal, provincial, and federal police forces. No, I will not give out my name just to satify your curiousity. Deal with it.

> *- misc.fitness.weights*

There can be real harm in being ``seen" online. One Usenet troublemaker forwards postings to their authors' supervisors, claiming that they were inappropriate uses of the net and that the author is a troll, etc. Although the original posts are completely legitimate (questions about integer precision in database packages and the like) many managers know little about Usenet culture and will assume that the employee must have been doing something wrong - and doing it on company time and under the company name. In an online discussion about this case, several people mentioned that, although they had free Internet access through work, they subscribed to a commercial service for their personal use, particularly for Usenet discussion: ``I'd rather pay an ISP to maintain a home account than risk getting some nut-case harassing my employers".

Privacy is a common reason for using a pseudonym, for Usenet is an exposed public forum in which the writers have no control over who reads their posts. People who are embarrassed use pseudonyms, such as system administrators who are asking how to fix something with which they ought to be familiar. People who are revealing extremely personal data (as in *alt.support.depression*) or who are discussing matters of dubious legality (as in *rec.drugs.psychedelic*) often use anonymous remailers. Finally, some people may simply not want their participation in Usenet, no matter how innocuous, to be public knowledge.

It is useful to distinguish between pseudonymity and pure anonymity. In the virtual world, many degrees of identification are possible. Full anonymity is one extreme of a continuum that runs from the totally anonymous to the thoroughly named. A pseudonym, though it may be untraceable to a real-world person, may have a well-established reputation in the virtual domain; a pseudonymous message may thus come with a wealth of contextual information about the sender. A purely anonymous message, on the other hand, stands alone.

Anonymity (including pseudonymity) is very controversial in the on-line world. On one side, anonymity is touted as the savior of personal freedom, necessary to ensure liberty in an era of increasingly sophisticated surveillance. It ``allows people to develop reputations based on the quality of their ideas, rather than their job, wealth, age, or status.'' [May 94]. On the other side, it is condemned it as an invitation to anarchy, providing cover for criminals from tax-evaders to terrorists. The ``very purpose of anonymity'', said Supreme Court Justice Scalia, is to ``facilitate wrong by eliminating accountability.'' [quoted in Froomkin 95].

There merit to both sides of the argument, much of it contingent on the distinction between anonymity and pseudonymity. Many of the strongest proponents of cryptographic privacy would agree that ``anonymous community'' is an oxymoron; their ideal is a pseudonymous world with merit-based reputations [May 94]. Purely anonymous individuals are capable of communicating with each other, but there is no accretion of personal histories in their interactions: reputation of any kind is impossible in a purely anonymous environment. The motivation for many of the qualities we associate with community, from cooperative behavior to creative endeavor, depends on the existence of distinct and persistent personas.

An interesting question is the accountability of a pseudonymous persona. The sanctions to offensive on-line behavior can be roughly divided into two main categories: those that involve making a connection to a real world person and those that do not. Complaints to a system administrator or other real-world authority are in the former category; killfiles and public castigation are in the latter. E-mail flames are somewhere in-between - one must know an electronic address that the offender accesses in order for them to be seen at all, but that address may be quite securely pseudonymous in relation to the real-world identity. In an electronic environment in which pseudonyms are prevalent, only the sanctions that do not require a connection to the real world are practical. While these mechanisms can only discourage, and not eliminate, outlawed behavior, they can have a significant effect [Kollock & Smith 95].

The evolving virtual world

In the world of biology, changes in signalling behavior may occur quite slowly, over evolutionary time. In the world of human interaction, changes can occur quite quickly. If excessive deception makes a signal lose its meaning, it can be replaced by a more reliable assessment signal or the community may begin to punish deception. In the virtual world, both the participants and the environment itself change: the participants establish new styles of interaction and the environment evolves as it is further designed and developed.

Killfiles are a good example of a social action that is poorly supported by the existing technology. One of the basic features (or drawbacks) of Usenet is that the readers are invisible. On the positive side, this lends it an aura of intimacy that would quite possibly be lacking if each writer were viscerally aware of the enormous number of people who follow the newsgroup. On the negative side, it makes the fact that one is ignoring someone very hard to indicate. The need to publicly turn away from someone can be seen in the custom of sending a posting that says ``Plonk!'', in response to that last-straw posting that caused one to killfile someone - ``plonk'' is the sound of dropping the offender into the killfile.

> Why do people feel the need to announce to the whole world that they have
> justplonked someone? Big deal. It happens all the time. Plonk 'em and get
> on with your life.

``On the Internet, no one can hear your killfile.'' Sometimes it is not only neccisary to regard someone as an idoit, you have to make them aware of it as well. I just wonder where they got the idea that killfiles have sound effects...

- alt.cypherpunks

The counterpart of ``plonk'' is the posting that simply says ``yes, I agree!''. These affirmations show that a particular opinion has enthusiastic backers and they provide a way to indicate an affiliation to an idea or person. These are frowned upon in Usenet etiquette because as a full-scale posting, they require too much time and effort to download, given the minimal information they include.

The online world is a wholly built environment. The architects of a virtual space - from the software designers to the site administrators -shape the community in a more profound way than do their real-world counterpart. People eat, sleep, and work in buildings; the buildings affect how happily they do these things. But the buildings do not completely control their perception of the world. In the electronic domain, the design of the environment is everything. Whether or not you know that other people are present or privy to a conversation, whether you can connect an on-line identity to a real-world person, whether you have only a faint notion of the personalities of those around you or a vibrant and detailed impression - this is all determined by the design of the environment.

How can Usenet - or other discussion-based systems - be redesigned to allow for better communication of social cues? Systems that are able to show participants or participant behavior - how other readers have navigated a newsgroup or how close or far other readers place themselves from an idea or person - are technically feasible. The real question is how they would effect the Usenet community. My prediction is that making the social patterns more visible would increase the strength of social pressures, making the community both more orderly and less spontaneous. But predicting the social ramifications of technology is difficult, especially when the whole environment is in constant flux.

For example, Usenet postings used to be ephemera, remaining available for only days or weeks before they disappeared from the net. Starting in 1995, several news archives have become available. These archives extend the lifetime of a posting indefinitely and, more significantly, they are searchable. One can request a listing of someone's entire Usenet oeuvre. Without such a search mechanism, finding all of someone's postings was nearly impossible: you might know that they were a frequent contributor to, say, the nutrition and medical groups, but have no idea that they spent their evenings as a verbal warrior in the ethnic disputes on *soc.culture.turkey* or writing baby-talk ``meow-chat'' postings to *rec.pets.cats*. The archives bring forth all of one's contributions for public examination, removed from the social context for which they were written. It involves a paradigm shift, from perceiving Usenet as a series of effectively private areas, bounded not by technical means but by their sheer numbers and parochial focus, to seeing it as a public repository of neatly cross-referenced postings.

This is not necessarily a harmful development for the Usenet community. One of the drawbacks of the virtual world has been that one's view of others is sketchy and one-sided. Being able to gather a more complex image of one's fellow participants can deepen the social ties as the users see each other as more fully-rounded individuals [Sproull & Kiesler 91]. Again, prediction is tricky. As awareness of the new paradigm increases, people may become far more concerned with managing their on-line reputation, resulting in widespread use of multiple pseudonyms - and an even murkier view of who's who on-line.

The Usenet reader's picture of the other participants is also being filled in by the Web. Whereas the archives present a documentary recording, the Web-based home page presents a crafted self-presentation [Donath & Robertson 94], showing how one wishes to appear - which can, of course, be quite inadvertently revealing [Goffman 59]. As home pages grow increasingly elaborate, their value to their creator grows. While it may not be terribly costly to discard, say, a name on AOL in order to escape from the consequences of actions done under it, one is far less inclined to abandon an online presence that has taken great effort to create. With an increasing number of articles signed with the writer's Web page, the Usenet readers gain both a deeper context for understanding an author's view, and a greater commitment by participants in the virtual environment.

Here again, the social ramifications may be unexpected. In a forum where a link to one's Web page is the norm, the opportunity to explore multiple persona's may be greatly curtailed (though perhaps given greater depth, if one was then motivated to create a elaborate series of pseudonymous portraits). For most people, one Web presence suffices - and it is often an official one, created for one's employer, one with a picture revealing age, race, gender, etc. The cost of deception would certainly be higher - the question remains whether that is necessarily a good thing.

New ways of establishing and of hiding identity are evolving in the virtual world. There is no formula that works best in all forums: balancing privacy and accountability, reliability and self-expression, security and accessibility requires a series of compromises and trade-offs whose value is very dependent on the goals of the group and of the individuals that comprise it. The role of this paper has been to examine closely the approaches - the signals - that have developed in a very diverse yet technically simple environment. What we have seen is a world of complex interactions, one that intermingles people from disparate real-world cultures and disparate virtual-world cultures; a world in which the boundaries exist only as social mechanisms and are both fluid and surprisingly durable. It is a world in which a technology built for the exchange of scientific data among a small class of academics and professionals has evolved into a communications forum in which information is still exchanged, but so is support and affiliation and adolescent bonding and outbursts of anger. It is a world that has evolved an intricate system of signals and behaviors that aid in establishing identity and in controlling identity deception.

(Draft 3.0)

Footnotes

[1] A FAQ (Frequently Asked Questions) is a document that contains many of the facts and anecdotes relevant to a group; their purpose is to answer these questions before they are asked - yet again - in the newsgroup. A FAQ may be quite long and require much an effort to create and keep up to date. [Return]

[2] Examples include "femme fatale" fireflies, brood parasites such as the cuckoo and the cowbird, and Batesian mimics such as the burrowing owl and the viceroy butterfly. "Femme fatale" fireflies are predatory females of the species Photuris who are able to mimic the flash pattern of females of the species Photinus. The deceptive Photuris female signals, the unsuspecting Photinus male approaches, and the predatory female attacks and eats him. Brood parasites lay their eggs in the nest

of another bird. The unwitting adop tive parent hatches the egg and raises the parasite, often at the expense of its own offspring. Batesian mim ics are harmless species that imitate species that are repellent to predators or competitors: the Viceroy butterfly resembles the bad-tasting Monarch; the hissing call of the burrowing owl sounds like a rattle snake's rattle. [Return]

[3] News is accessible by anyone with Internet access. How many people that is at a given moment is debate able, but it is agreed that the number is growing exponentially. (Oct 1994 estimate 13.5 million: should update this number as close to publication as possible - could be 10x or more this number by 1997). [Return]

[4] Names and other identifying features have been changed. [Return]

[5] Killfiles are used to filter out postings by people or about topics one does not wish to read. For more about killfiles see Identity concealment. [Return]

[6] This is not to say that these "parts" are necessarily straightforward. See O'Brien 97 for a discussion of the complexity of gender roles - and the added intricacy of their virtual manifestation. [Return]

[7] There exists a technological solution to this problem. A digital signature can ensure that a message has not been altered since it was signed and, given various levels of certification, it can guarantee that a particular person was the signer. Interestingly, the certification of identity is personal trust. Individuals vouch for indi viduals and their personal guarantees become a part of one's digital signature. If I know nothing about the people who vouched for you, the guarantee is meaningless.). They are currently rather difficult to use, though this is a problem more of interface than underlying technology. As encryption and decryption become an integrated part of the virtual environment, the appearance of a real, vouched-for persona may begin to differ markedly from other, more ephemeral beings. (See [Garfinkel 95] for a full technical exposition). [Return]

[8] A letter posted to Usenet is distributed through the net by being passed from the sender's machine through a series of Usenet sites, each of which distributes it to a number of other sites. When it finally reaches a particular reader's machine, it may have passed through 20 or more sites or "hops", each of which records its name on the header of the posting. While the exact route a posting from A-B will take is not predictable (one of the distinctive technological features of the Internet is its ability to re-route itself around down machine and clogged regions), obvious peculiarities in the route are signs of a forged message. [Return]

References and further reading

On identity, society, virtual community

Aronson 95
> Aronson, E. 1995. The Social Animal (7th edition). New York: W.H.Freeman and Co.

Beniger 87
> Beniger, J.R. 1987. ``Personalization of mass media and the growth of pseudo community'', Communication Research 14, no. 3:252-371.

Constant et al. 95
> Constant, D., Kiesler S. and Sproull, L. 1996. ``The kindness of strangers: On the usefulness of weak ties for technical advice,'' Organization Science 7, 119-135.

Curtis 92

Curtis, P. 1992. Mudding: social phenomena in text-based virtual realities. Proceedings of the 1992 Conference on Directions and Implications of Advanced Computing. Berkeley, May 1992.

Davis 92

Davis, F. 1992. Fashion, Culture and Identity. Chicago, IL: University of Chicago Press.

Dibbell 93

Dibbell, J. 1993. A rape in cyberspace. The Village Voice. Dec. 21, 1993.

Donath & Robertson 94

Donath, J. and Robertson, N. 1994. The Sociable Web Proceedings of the Second International WWW Conference. Chicago, IL, October 1994.

Donath 95

Donath, J. 1995. Sociable Information Spaces. Proceedings of the Second IEEE International Workshop on Community Networking, Princeton, NJ, June, 1995.

Fiske 89

Fiske, J. 1989. Understanding Popular Culture. London & New York: Routledge.

Froomkin 95

Froomkin, A.M. 1995. Anonymity and its enemies. Journal of Online Law, art. 4.

Garfinkel 95

Garfinkel, S. 1995. PGP: Pretty Good Privacy. Sebastopol, CA: O'Reilly & Associates

Geertz 73

Geertz, C. 1973. The Interpretation of Cultures. Basic Books.

Goffman 59

Goffman, E. 1959. The Presentation of Self in Everyday Life. New York: Doubleday.

Goffman 67

Goffman, E. 1967. Interaction Ritual. New York: Pantheon Books.

Kollock 94 v

Kollock, P. The emergence of exchange structures: an experimental study of uncertainty, commitment and trust. American Journal of Sociology. 100. 313-45.

Kollock & Smith 95

Kollock, P. and Smith M. 1995. Managing the Virtual Commons: Cooperation and Conflict in Computer Communities. Forthcoming in (S. Herring ed.) Computer-Mediated Communication. Amsterdam: John Benjamins.

May 94

May, T. 1994. The cyphernomicon: cypherpunks FAQ and more. http://www.oberlin.edu/%7Ebrchkind/cyphernomicon/

McCracken 88

McCracken, G. 1988. Culture and Consumption: New Approaches to the Symbolic Character of Consumer Goods and Activities. Bloomington: Indiana University Press.

Milgram 77

Milgram, S. 1977. The Individual in a Social World. Reading, MA: Addison-Wesley.Simmel, G. 1971. On Individuality and Social Forms. (D. Levine, ed). Chicago: The University of

Chicago Press.

O'Brien 97
> O'Brien, J. 1997 Gender on (the) line: an erasable institution? In Communities in Cyberspace. Berkeley, CA: The University of California Press.

Reid 91
> Reid, E. 1991. Electropolis: Communication and community on internet relay chat. Thesis, Dept. of History, University of Melbourne.

Rheingold 93
> Rheingold, H. 1993. The Virtual Community: Homesteading on the Electronic Frontier. MA: Addison-Wesley Pub. Co.

Saville-Troike 82
> Saville-Troike, M. 1982. The Ethnography of Communication. London: Basil Blackwell.

Simmel 71
> Simmel, G. 1971. On Individuality and Social Forms. (D. Levine, ed). Chicago: The University of Chicago Press.

Sproull & Kiesler 91
> Sproull, L. & Kiesler, S. 1991. Connections: New Ways of Working in the Networked Organization. Cambridge: MIT Press.

Sproull & Faraj 93
> Sproull, L. & Faraj, S. 1993. Atheism, Sex, and Databases: The Net as a Social Technology. Forthcoming in (B. Kahin and J. Keller, eds.) Public Access to the Internet. Prentice-Hall.

Stone 92a
> Stone, A.R. 1992a. Will the real body please stand up?: Boundary stories about virtual cultures. In M. Benedikt (ed.) Cyberspace: First Steps. Cambridge MA: MIT Press.

Synnott 93
> Synnott, A. 1993. The body social: symbolism, self, and society. London: Routledge.

Turkle 95
> Turkle, S. 1995. Life on the Screen: Identity in the Age of the Internet. New York: Simon & Schuster.

Wellman 97
> Wellman B. and Gulia M. Net surfers don't ride alone: virtual communities as communities. In Communities in Cyberspace. Berkeley: University of California Press.

On signals and modeling deception

Alcock 89
> Alcock, John. 1989. Animal Behavior: An Evolutionary Approach. (4th ed.). Sinauer Associates, Inc., Sunderland, MA.

Dawkins & Guilford 91
> Dawkins, M.S. & Guilford, T. 1991. The corruption of honest signalling. Anim. Behav. 41. 865-873.

Dawkins & Krebs 78

Dawkins, R. & Krebs, J.R. 1978. Animal signals: information or manipulation? in J.R. Krebs and N. B. Davies (eds.) Behavioural Ecology: An Evolutionary Approach. pp. 282-313. Blackwell Scientific Publications, Oxford.

Grafen 90b

Grafen A. 1990b. Biological signals as handicaps. J. Theor. Biol. 144. 517-546.

Hauser 92

Hauser, M. 1992. Costs of deception: Cheaters are punished in rhesus monkeys. Proc. Natl. Acad. Sci. USA. 89. 12137-12139.

Hauser 96

Hauser, M. 1996. The Evolution of Communication. Cambridge, MA: MIT Press.

Johnstone & Grafen 91

Johnstone, R.A. & Grafen, A. 1991. Error-prone signalling. Proc. R. Soc. Lond. B 248, 229-233.

Krebs & Davies 93

Krebs, J.R. & Davies, N.B. 1993. An Introduction to Behavioural Ecology. 3rd Edition. Oxford: Blackwell Scientific Publications.

Maynard Smith 74

Maynard Smith, J. 1974. The theory of games and the evolution of animal conflicts. J. Theor. Biol. 47. 209-221.

Zahavi 77

Zahavi, A. 1977. The cost of honesty (further remarks on the handicap principle). J. Theor. Biol. 67, 603-605.

Zahavi 93b

Zahavi, A. 1993b.The fallacy of conventional signalling. The Royal Society Philosophical Transaction B. 340. 227-230.

Last Modified: 02:16pm EST, November 12, 1996