Online communities: Usability, sociabilty, theory and methods

Jenny Preece
Chair & Professor, Information Systems Department
University of Maryland Baltimore County
1001 Hilltop Circle, Baltimore, MD 21250
Preece@umbc.edu, http://www.ifsm.umbc.edu/~preece

(Note: This chapter is adapted from (Preece, 2000) "Thriving Online Communities: Usability and Sociability", published by John Wiley & Sons. Permission from the publishers should be sought. Parts have also appeared in Brown et al. 1999)

Internet usage doubles every fifty-two days. Over eighteen million people are AOL subscribers– many of whom are attractive by AOL's emphasis on email, chats, instant messaging and, of course, the web. Active Worlds, a graphical chat environment, has over a million participants. During the first quarter of 1998 450,000 messages were posted to 20,000 UseNet groups. And the numbers of people coming online continues to increase as e-commerce, online education, online health, and increasing amounts of information and people with whom to chat, entices even more people. Meanwhile, according to Moore's computing power doubles every eighteen months. For the well educated, with jobs and Internet stock, the future looks bright. However, the gap continues to broaden between low and high income, poorly and well-educated people. There are fears that socializing face-to-face will decline leading to an unprecedented number of lonely, psychologically impoverished people.

Online communities is an emerging discipline in university curricula and research. However it is finding a place in agendas of major laboratories and national funding bodies (Brown, 1999b; Brown, 1999c). Like other new topics, its status is debatable. At a similar stage, researchers tried to characterize human-computer interaction in 90s. At an National Science Foundation workshop, Stu Card identified four stages in the growth of a discipline (Card, 1991) (Olson & Olson, 1997). Building and evaluating individual systems or *point systems*, is the first stage in early development. As more communities develop, research intensifies and *comparative evaluation* studies start to identify dimensions for success. This is the second developmental stages. The third stage is characterized by understanding the *relationships* between various developmental stages more thoroughly, so that *models*, *laws* and *theory* can be articulated – an indicator of the discipline reaching the fourth stage and coming of age.

Current research in online communities deals primarily with individual communities - i.e., stage one. However, critics argue that the basic technology (i.e., email, listservers, bulletin boards, chats etc.) are established technologies, which people use enthusiastically, so they question why research is needed. There are four reasons:

- The Internet enables millions of people to communicate online at once. Little is known about the social and technical dynamics of mass communication involving thousands and millions of participants via the Internet.

- Online populations will be diverse in terms of culture, age, technical and educational experience. Technology with appropriate usability for this broad range of users is needed.

- People have expectations of excellent usability and sociability. At present these expectations are not being met.

- Local, national and international agencies, governments and e-commerce will demand that people go online for certain activities such as voting, paying taxes, licensing, social support, health, all kinds of information and purchasing goods. Online communities *must* be well developed so that they are usable by *all* citizens.

Directions[1] for research and development are needed that address usability and sociability problems so that better online communities can be developed. There is a particularly strong need to involve social scientists as well as computer scientists. Successful online communities will result from a blend of good usability and carefully guided social policies, that encourage social interaction. Theory and better research methods are also needed to support Internet research and system development.

Usability and Sociability: A Framework for online community design

There is no formula for thriving online community. Online communities are new and there is a dearth of experience to guide successful development. Furthermore, online communities evolve and change constantly depending on their membership. What may be important early in the life of a community may not be significant later on. Success is determined by three key factors: *usability*, *sociability* and their affect on the *interactions of community members*. Developers have *little* or *no* control over community members, except in some e-commerce communities where interaction is strongly managed. However, developers can do much to set the tone of a community by designing or selecting software with good usability and developing suitable sociability.

Software with good *usability* supports *rapid learning*, *high skill retention*, *low error rates* and *high productivity*. It is *consistent*, *controllable* and *predictable* making it *pleasant* and *effective* to use. Usability is a key ingredient for the success of *any* software. Good usability supports people's creativity, improves their productivity and makes them feel good. Poor usability leads to frustration, wastes time, energy and money. Norman describes the transactions between humans and computers in terms of crossing two gulfs (Norman, 1986). The first, is the *gulf of execution* and the second is the *gulf of evaluation*. Meaningful commands, menus and icons, a well-designed navigation system, and comprehensible messages help to reduce the cognitive and physical effort required to bridge these gulfs. Graphical user interfaces (GUIs), based on appropriate visual metaphors, for example, help by enabling users to infer knowledge about the target computer system from their knowledge of the metaphoric system. Three important principles of usability are that systems should be consistent, controllable and predictable (Shneiderman, 1999).

Consistent software uses the same terms and procedures for achieving the same functionality throughout the program. The notion of consistency is far reaching. Sequences of actions should follow the same format. Color, typography and terminology should also be consistent and so should. Users want software that supports but does not take away their sense of *control*, so they can do what they want, when they want and not be constrained by the software. Software that is consistent and controllable is predictable too. *Predictable* software enables users to continually build on their experience so that they build confidence and skills with experience.

There is a large collection of research and many books provide guidelines for user interface design (e.g., (Nielsen & Mack, 1994) (Preece et al., 1994; Shneiderman, 1999) (Hackos & Redish, 1998)). An increasing number focus specifically on web design (Lynch & Horton, 1999) (Spool, Scanlon, Schroeder, Snyder, & DeAngelo, 1997). However, although there are excellent books that deal with community development (Schuler, 1996) (Figallo, 1998) few pay much attention to usability, but as more people gain

---

[1] European Union and National Science Foundation report, June 1-4, 1999, prepared by Christoph Busch, Fraunhofer IGD, Germany; Richard Guedj, INT, France; Wendy Kellogg, IBM T. J. Watson Research Center, USA; David Leevers, VERS Associates, UK; Sudhir Mudur, National Center for Software Technology, India; Jennifer Preece (lead author), University of Maryland Baltimore County, USA; Ben Shneiderman, University of Maryland College Park, USA; John Thomas, IBM T. J. Watson Research Center, USA; Deb Roy, MIT, USA; and Junji Yamaguchi, Independent, Japan.

experience of online communities, there will be cries for better usability (Gaines, Chen, & Shaw, 1997; Preece, 1999a).

Developers *cannot* control social interaction but careful, early social planing, coupled with good usability encourage thriving communities to evolve. *Sociability* focuses on *social interaction*. Communities with good *sociability* have *social policies* that support the community's *purpose* and are *understandable*, socially *acceptable* and *practicable* (Preece, 2000). Unlike usability, the concept of sociability is not well understood. There are many questions to address including the following. How does a community communicate its purpose? What are the impacts of different personalities and policies (e.g. joining regulations, controlling anti-social behavior, keeping conversations on track etc.)? How should emotion, content and online activity be represented? How can privacy and security be assured? In this chapter guidelines for sociability are presented. Populations of communities range from millions to small communities of less than twenty. There is also huge variation in posting rates (Nonnecke & Preece, 2000). Some people interact frequently others may post just once every few months. Some people lurk and never send messages. Perhaps with different software or a smaller, safer-feeling environment, lurkers might post? Potentially, this large volume of lurkers represents huge revenue for e-commerce.

The relationship between usability and sociability in design of online communities is important for their success. As in any software development it is essential to have a detailed understanding of users' needs (i.e. requirements), and these will determine functionality and usability. However, software design alone does not determine success. While online community developers cannot control social interaction, they can strongly influence it by the policies that they put in place and how they manage these policies. With careful planning, well-designed policies can help the community to develop. In order to do this, however, a much better understanding of social interaction online is needed. For reason it is essential that social scientists be involved in future research and that there work informs software design. As the arrows in figure one show, online communities are constantly evolving and the community's needs have to be revisited and reflected in software design and usability. Furthermore, usability ad sociability are strongly linked, and this emphasizes the need for social and computer scientists to collaborate.

Community membership registration is an example of the relationship between usability and sociability. Decisions about whether to have registration and what the policy should involve are sociability considerations. For example, making people register will deter frivolous posting and flaming and help to eliminate spam. However, it might also deter many people from joining who would do so, if the community was open. Social planning (i.e. sociability) involves weighing-up these decisions and deciding what is best for the community depending on its purpose and its membership. Such decisions cannot be divorced from usabilty. For example, decisions about whether to have a registration policy and the nature of that policy are sociability decisions. The actual form, interactive procedure, positioning and wording of the policy are usability issues.

Research topics for usability and sociability of online communities are discussed in the next two sections. (Preece, 1999b). Many of the ideas presented are from an earlier paper submitted to the EC/NSF joint Research Workshop held at Bonas, France in June 1999. However, as a result of contributions from members[1] of the workshop, this report was improved and published as part of a report of the full workshop, versions of which have been published widely (Brown, 1999b; Brown, 1999a; Brown, 1999c). A further round of changes was made when the original report was conceived within the usability and sociability online community framework (Preece, 1999b). The current version presented in the next two sections represents yet another iteration.

Usability

Three areas of central importance for online community development are: design and representation; security and privacy; and scalability.

Design and representation

Knowledge and theories about culture and social activity in online communities provide a foundation for software design. Research is needed to develop representations to reveal online behavior as it happens,

histories of behavior, stored communication and knowledge, nature of communication (e.g. which topics were discussed), the number of people participating and relationships between participants. In addition, individuals may want to represent themselves in different ways using avatars and other representations.

- *Revealing behavior*
  What is the impact of revealing behavior of individuals and groups in different circumstances? How can online communities support different kinds of behaviors and information, e.g., self-expression, humor, personality, mood, identity, empathy, aesthetics, age etc. What is the impact of different types of representation for community dynamics – e.g. people coming, going and engaging in different behavior? What happens when dealing with different modalities?

- *Interactions between representations and social processes*
  What are the interactions between different representations and social processes? What kind of new communication processes might arise from use of different representations? Research indicates that socio-emotional processes are different online, particularly in textual environments which do not cues from body language and voice tone.

- *Revealing the content and emotion of messages*
  How should different content be represented. For example, researchers have demonstrated that computer mediated communication facilitates communication about sensitive, personal or emotional subjects distinguishing between emotional content, empathy and factual information. Similarly, how can people reveal their mood. Words and phrases are available at the press of a function key in some systems. Emoticons are available in others. Avatars can sometimes become animated so that they wave or jump. What other ways might there be and what is the affect of enabling people to reveal both content and emotions explicitly? If a collective mood of a group could be established by, for example, a voting system, how could it be expressed and would this encourage better communication? How can users understand the mood and nuances of community when they come online? Similarly, when newcomers join the community what mechanisms would help them to guage the feelings of the group, depth of conversation, duration of discussing a single topic and so on?

- *How are large communities represented on-line?*
  Large numbers of people wanting to join online communities present new challenges for designers. What types of features should be provided in software to support large communities, their moderators and administrators? In chapter 3 some examples of representing changing dynamics in small chat communities were discussed, but little attention has been given to scaling such approaches to represent thousands of participants. How, for example, do you represent a million people in a community and what should the interface be like?

- *Weight of issues*
  What characteristics of online environments bias for or against different aspects of communication? How do they get balanced and traded off? For example, what kinds of interfaces such as visualizations are needed to support individual differences among users?

Security and privacy

Success of some online communities will be strongly influenced by how secure they are. Personal health and credit card details must be secure. Systems must not only be secure, users must trust the security. Incorporation of strong cryptographic protocols is essential. These protocols realize classical security requirements such as mutual authentication of communication between trading partners, confidentiality of the transaction, authenticity and integrity of the goods. In addition the availability and integration of adequate payment protocols is essential to satisfy the needs of commerce. Security is a major technical issue for many online communities. Identifying what security is needed and how to present security procedures to users are usability and sociability concerns.

Two crucial areas require substantial research in the near future in order to develop a powerful electronic market and eliminate lack of trust in online transaction:

- *Conflict between identification vs. privacy or anonymity*
  There is a conflict of interest between vendors and content providers in e-commerce systems, which demand identification of their customers using concepts such as globally unique identifiers to realize digital fingerprints (attached to digital goods). In contrast there is the crucial demand for privacy protection raised by consumers, that requires non-traceable interactions. Research is required to overcome this conflict, which might be resolved using limited pseudonymity and set up of Trusted Third Parties.

- *Copyright protection*
  As content contributed to online systems in commercial and non-commercial instances grows, the question of intellectual property protection gets more and more crucial. Consequently, adequate copyright protection mechanisms must be investigated and developed. In this context digital watermarking is a promising technology to provide security to content contributors, since it allows an imperceptible (inaudible) mark to be embedded in the content data itself, that bears the identity of the copyright holder and, eventually, also the purchaser's fingerprint. Research must focus on watermarking technology for all kinds of multimedia data, namely images, audio, video and 3D data such as representations of virtual models and virtual environments. While substantial work in the protection of images is already underway, protection of virtual worlds (i.e. the 3D model itself) is lagging far behind.

Scalability

Scalability is a research priority with strong technical challenges and implications for usability and sociability. With so many people from across the world wanting to join or develop online communities, research is needed to produce software with high usability and guide sociability in very large communities. For example, how should thousands of people be represented online? How should conversations be structured? How do we guide online crowds and develop social protocols? What kind of online governance is needed to control hostile behavior, spamming, sexually explicit behavior, and ensure voting practices are fair? Research is needed to develop interfaces to support large populations of participants, and tool development for moderators.

Sociability

Communities with good sociability have social policies that support the community's purpose and are understandable, socially acceptable and practicable. In order to develop communities with good sociability a much better understanding of community, culture and social interaction online is needed. Strong multi-disciplinary research teams are needed with social scientists leading basic research initiatives. The role of computer scientists will be to transfer this knowledge into software that supports social interactions, protects individuals' privacy, provides security and encourages universal access. Important research topics include: community and culture; and ethical issues and universal access.

Community and culture

We need to understand cultural differences better and how to support diversity online. Communities of all kinds are rich social environments, which cannot be observed through the port holes of individual human-computer interfaces. Many communities have complex lifecycles punctuated by temporal events, such as life-cycle phenomena and unanticipated events. So far work in human computer interaction and virtual environments have not focused on effective online community design and maintenance. Solutions to most of the questions discussed in this section will support better sociability planning.

- *Understanding difference between communities*
  Some research issues that need to be investigated include understanding differences between networked communities, virtual communities and virtual environments – i.e. the full range of the physical and online spectrum of communities. How do the social policies and structures adopted by these different communities compare, differ, and relate to each other? What types of governance procedures do they need?

- *Life cycles*
Physical life has daily cycles, that are often upheld by eating rituals, breakfast, lunch, dinner, work, periods of relaxation and sleep. Time zone differences cause severe practical limitations for synchronous communication. Weekly and annual cycles vary from culture to culture, time zone to time zone. What is the affect of these cycles on international communities? Real communities also have generation cycles - birth, teenage life, marriage, parenting, retirement, old age and death. Perhaps online community developers can take advantage of humans' propensity to structure behaviors around cycles. TV producers appear to understand these cycles well. They can predict maximum viewing times for different sectors of a population. How might software support these lifecycles? What role do rituals have off-line and how might they translate online? For example, some communities have experimented with synchronizing snacks to enforce a sense of community. At an appointed time, participants distributed across several locations will have an M&M or cookie and soda break. However, anecdotes from colleagues that have participated in such events, suggest that they don't contribute to the sense of community. There is no substitute for being there when it comes to sharing a nice meal or a hug!

- *Interaction dynamics*
Interaction dynamics are different in four-person peer discussion, a twenty five person class room, a group of one hundred and fifty friends, and several thousand people meeting for a conference. Each has there own social rituals. In the physical world participants have different expectations in these different social settings. Physical space and the number of people present provide strong cues, which are not available online. The mere act of going to a meeting often conveys a strong message of commitment. The person has risked traveling, invested time and money in being physically present. Handshaking, hugging and eating are social activities that don't translate well online. Augmenting our senses with wearable computers, empowering creativity with well-designed tools, enabling people to indicate their emotional states etc. could improve the quality of interaction. Access to information, stories and monitored data could be useful to both individuals and communities for developing greater collective intelligence.

Ethical issues and universal access

Designing software that supports good ethics and universal access is a challenge for technical and social scientists. Basic sociological research is needed to development of codes of conduct online and other forms of sociability. However, in order to make the features available to support universal access, challenging computer science issues need solving. Good usability design will be needed to make new functionality usable by the wide range of users for whom it is designed.

- *Ethical issues*
Awareness of the dangers of participating online
Many users fail to appreciate the potential dangers of participating online. They are unaware that correspondence is electronically stored and of when they initiate insecure communications or operations. Current software does not give adequate feedback about such operations. Dialog boxes that users generally disable because they are annoying and tiny icons provide insufficient warning. Better ways are needed of protecting users by notifying them about unencrypted or encrypted traffic, potential persistent storage in which their privacy is endangered, and acts with possible legal consequences such as digital signatures.

Codes of conduct for online communities
Protection of individual and group privacy in online communities can be partially realized through improved technology. In addition policies are needed to ensure host operators and maintainers are required to follow fundamental privacy rules. As an analog to the ethical binding of physicians a Hippocratic oath could be defined for operators and key persons of online communities. Research is also needed to identify successful models of self-governance in online communities. A better understanding is needed of how co-operation, trust and empathy develop online, how these relationships change over time and how changes in population size and demography affect them. Contributions from social and political scientists will be needed to develop different models of governance and integrate online policies with pre-existing local, national and international policies.

Improved environments
Research is needed to understand the impact of digital technology on cultural diversity, environmental issues, conservation of limited resources and changes in people's standard of living. This information will help to inform national and international agencies so they can deploy resources well.

- *Universal Access*

A report entitled "Falling through the digital divide" (National Telecommunications and Information Administration & Commerce., 1999) produced by the US Department of Commerce shows that although the gap between the number of men and women online is diminishing, the gap between rich and poor, well-educated and poorly educated is increasing. Lack of access to computing equipment affects both individuals and nations. To ensure universal access to online communities for people of all ages, cultures, languages, income levels, educational, physical and mental abilities, five broad areas of research are identified. The first, multiple interaction modalities, calls for alternatives to text input and output. The second area suggests research into adaptive interfaces that can be tuned to a wide range of communication abilities and preferences. Third, research is needed into technologies for supporting interaction in any language opening the door to communities with 'around the world access'. Fourth, translation technologies are needed for bringing together different language groups.

Multiple interaction modalities
Spoken language input accommodates individuals who are unable to type due to illiteracy or motor control impairments. Speech interfaces are also useful for individuals operating in disabling situations where hands and eyes are busy with other activities. Spoken language interaction does not necessarily imply the use of automatic speech recognition (ASR) or text-to-speech synthesis (TTS). Alternatively, speech may be represented simply as acoustic waveforms without conversion to and from text. Each representation raises a different set of research questions. Recent advances in computer vision make possible human-computer interfaces based on face and body gestures. Gestures may be a natural modality for conveying non-linguistic information including a person's mood and various conversational cures which express willingness and interest to engage in dialog. Gestures may also provide a linguistic input modality for individuals unable to type or speak. Touch tablets and buttons may also be used for input, for example for individuals with impaired speech and fine-level motor control.

Individual needs and preferences
Interfaces are needed that are easy to learn and tailorable to the needs of individual users. Individual differences to take into account include: education, preferences for using different applications, age, language and culture. Access to equipment will also be an issue, interfaces that can be tailored for low-end technology and slow modems as well as satisfying state-of- the art users. Different approaches need to be explored including tailorable interfaces, user modeling and adaptive interfaces.

Multilingual support
Digital communications hardware and software is designed primarily for English text. Future systems must support a variety of fonts and keyboard mappings for various languages. Keyboard maps provide challenges for languages which have more characters than keys in the standard 101 key board making a s one-to-one mapping impossible. Similarly, many languages do not follow the strict left-to-right conventions of English that must be considered for text mapping. Difficult research issues need to be addressed for supporting multilingual spoken language interaction. Currently speech recognition and synthesis technologies exist for less than twenty languages. These technologies are extremely costly to develop for new languages. New development strategies are needed for creating speech technologies for multiple languages that require less development effort, perhaps at the cost of loss of some performance. Technologies for large numbers of languages are needed which may be refined over time. Universal access should also include multilingual spoken language interaction since the majority of the world's population in developing countries is illiterate.

Inter-lingual support
A grand research challenge is to design systems that support communication between people who do not share a common language. In addition to machine translation of text and speech, we should also investigate

visual languages and use of multimedia. Tools that enable communities to dynamically create shared resources, including multilingual, multimedia dictionaries and thesauri to support common codes of communication, are needed.

Theory and methods

Many potential candidate theories exist but most address a small area and despite being grandly called theories, they lack the breadth, explanatory and predictive powers. Online communities research needs high-level, generalizable theories that enable us to:

- *understand* communication in different types of communities;
- make *predictions*; and
- *inform* online community design.

There are three general categories of theories address:

- one-to-one or small group communication via different media;
- social interaction and community networks; and
- relationship between software design and social behavior.

One-to-one and small group communication via different media

Social presence theory (Short, Williams, & Christie, 1976) (Rice, 1987), *media richness theory* (Daft & Lengel, 1986) and *social contextual cues theory* (Sproull & Kiesler, 1991) differ in detail, but all broadly assert that reduced social cues resulting from low bandwidth environments encourage impersonal, uninhibited and less social behavior. *Social presence,* the most widely accepted of the three, concerns how well different media convey a sense of presence to communicating partners (Short et al., 1976). Social presence depends not only on the words but also on verbal and non-verbal cues and the communication context (Rice, 1993). Communication media differ in how much they can overcome constraints associated with lack of social presence (Rice, 1993). Understanding the relationship between the social presence of different media and the needs of the communication task is vitally important for online community developers.

Common ground theory provides a framework for understanding how two people ensure they are communicating with each other through a process of grounding. How grounding occurs varies from situation to situation. Grounding takes one form in face-to-face conversation and another in computer mediated communication. It takes one form when calling directory assistance, another when chatting with a friend and yet another in intellectual debate. Grounding is, therefore, influenced both by the communication medium and the communication task.

Clark and Brennan ((Clark & Brennan, 1993) p. 229) identify *constraints* that a medium may place on common ground such as: co-presence (i.e. perception of sharing same physical environment), visibility (participants can see each other), audibility, co-temporality, reviewability, revisability etc. Overcoming the constraints of a medium imposes a cost that places a burden on grounding. Although the term 'cost' implies a disadvantage, there are occasions when a cost can confer an advantage. For example, 'delay costs' may be a problem in a video conferencing environment where communicators expect turn taking as in face-to-face communication. In textual conferencing however, delays can be valuable because they give time for reflection. In a study of recovering alcoholics in a listserver community, several participants commented that they liked having time to reflect and compose (King, 1994).

Online chats appear to be severely limited. Not only can participants not see each other, they can send only short remarks which they must generate quickly in order to catch the flow of the conversation. Furthermore, there is little time for reflection, no time for fault correction and speaker change-over is chaotic. However, people who use chats regularly are enthusiastic and do not have these complaints. This indicates a possible weakness of the theory is that it does not take account of the communication task, only the media is important. Clearly certain media support particular communication tasks better than others.

Focusing on media without regard to the task for which they are being used will give a false impression of their efficacy.

Other theories that help to explain human behavior and emotions, such as empathy (Levenson & Ruef, 1992) would also be useful for informing design (Preece, 1998; Preece, 1999a). Understanding development of trust online and providing usability and sociability support has become important in e-commerce. Little work has been done in this area and much more research is needed. Trust is the expectation that arises within a community of regular, honest and cooperative behavior, based on commonly shared norms and part experiences (Fukuyama, 1995 p. x). Trust is also important for any community concerned with knowledge exchange and management (Liebowitz, 1999), particularly if sensitive information is involved. An added complication for any interaction involving software is that the software itself may be designed to suggest trust because systems designers impart social and moral values in their work (Friedman, 1997 p. 1), which can be dangerous if users are deceived into falsely developing trust.

While these theories help to explain why flaming and other forms of behavior may occur, they speak about dyads and small groups not communities. How such theories as common ground scale up is debatable. Recently Whittaker and his colleagues (Whittaker, Terveen, Hill, & Cherny, 1998) investigated common ground between participants in Usenet groups. In this study 2.15 million messages were analyzed from 659,450 posters, collected from 500 newsgroups over a six-month period. Although common ground was inferred from the threading patterns and used to justify the claim that cross-posting between lists may be valuable in stimulating conversation, the efficacy of this approach is far from clear.

Community networks

Many theories from sociology, communication studies and psychology can be imported to help explain and predict behavior in online communities. However, none present a coherent leading theoretical approach. A foundational unifying theory that *predicts* online behavior and *guides* usability and sociability design would be helpful.

Critical mass theory claims that a critical mass of participants is needed to make online participation worthwhile (Markus, 1987) (Markus, 1990; Morris & Ogan, 1996) (Rice, 1994) (Ackerman & Starr,). However, it provides no indication of what this magic number might be under certain circumstances for different kinds of communities, so its value to designers is limited.

Social network analysis (Wellman, 1997) (Wellman, 1992) (Haythornthwaite, 1996). is a promising theory and set of techniques from sociology. A *social network* is a set of people (or organizations or other social entities) connected by a set of social relations, such as friendship, co-working, or information exchange (Garton, Haythornthwaite, & Wellman, 1999 p. 75). Computers support social networks by linking people who are geographically dispersed who otherwise could not relate to each other easily. Social network analysts seek to describe networks of relations as fully as possible, tease out the prominent patters in such networks, trace the flow of information (and other resources) through them, and discover what effects these relations and networks have on people and organizations. Analysts look for resources that are exchanged in creating and maintaining social relationships. These relationships can be strong or weak, uni- or bi-directional and occur over long or short periods of time. Pairs who maintain *strong* ties are likely to share what resources they have, even though, for some people these resources may be severely limited. In contrast *weakly* tied people are less likely to share resources. However having many weak ties tends to provide access to more diverse types of resources because each person operates in different social networks and has access to different resources (Garton et al., 1999).

Successful communities are built on *co-operation*. If community members do not feel a sense of responsibility towards each other they are unlikely to co-operate. Unfortunately there are often times when behavior that seems perfectly reasonable to an individual and gets the person what they want is damaging to the group. This tension between the group and the individual is known as a *social dilemma* (Kollock, 1998). Social dilemmas form the basis of much research on inter-personal co-operation in psychology, sociology and anthropology. One of the most well-known studies is Axelrod's two-person situation called prisoner's dilemma (Axelrod, 1984), in which three conditions for co-operation are identified.

The first condition is that it must be likely that two individuals will meet again in the future, otherwise there will be a temptation for one person to get what he or she wants and not worry about the other person because there are no implications from further interactions. This suggests that online communities need to foster on-going relationships. How this can this be achieved? People can drop into a Usenet group, send a message, and disappear never to be heard of again. There are no consequences for the person behaving this way but the effects on the community may be devastating. A registration procedure is one way for communities to deter casual hopping from community to community, but this can have other drawbacks, such as some people may be deterred from joining.

The second condition is that people should be able to identify each other. This condition has far-reaching implications for online communities, particularly those that use textual media where seeing facial expressions is not possible. Observing a person's presence in the environment from their message provides some clues, especially if the group is small and people already know each other. However, this is not the case in Usenet and other large open groups. Other identifying characteristics include login-in identification and chosen names, but in textual environments these can be woefully inadequate. Avatars could be a step in the right direction but while they provide an identifiable character, they hide the real person. Being able to reliably identify someone online is a complex issue. There are good reasons why people want to be anonymous. Unwanted soliciting either to sell commodities or for romantic encounters, or to obtain information about a person, such as medical records or financial information, is a real threat. Whether the Internet should be controlled or even policed is a difficult and contentious issue, which raises concerns about free speech. Not being able to identify people deters trust and encourages unsociable behavior. However, communities have different goals and they differ hugely, so probably different solutions are needed to deal with these problems. Further research is needed to open up understanding of these issues.

The third condition is the need for information about a person's past behavior, which means that some kind of history-keeping mechanism is necessary. Since most people communicate with many different people online and many online communities are getting very large, it is difficult to remember people, particularly without physical cues. Those systems that keep archives often do not have good search facilities. In any case, searching takes time. It is hard to see how this problem could be ameliorated. Perhaps a small history of recent interactions and major incidents could be generated each time a person sends a message. Reputation managers such as the one in E-bay may also be a way forward.

What is acceptable in any community will depend on the aims of the community, the people who contribute and any rules that have been agreed. The type of community can have a large impact. Professional and scholarly communities focus on communicating ideas and information, in a socially acceptable way. They can be controversial. Political communities are concerned with persuasion and debates that can become emotionally charged. Determining the boundaries between acceptable and unacceptable social behavior involves distinguishing between debate about ideas, which is acceptable, and *ad hominem* attacks directed at specific people which should not be tolerated. Support communities have a lower threshold for aggressive behavior, because it can be damaging to people already dealing with problems.

Research is needed to develop suitable theories of mass social interaction that will predict and guide online community development.

Influence of software design on social behavior

The last area of theory that would be helpful for online communities' involves better understanding of how software design can impact users' perceptions and behavior. Several already existing theories may shed light on this issue. For example, s*ituated action* explains how people's use of technology is influenced by the environment in which they are situated (Suchman, 1987). *Adaptive structuration* is concerned with the influence of software design on task execution and performance, and how people adapt their behavior to make best use of the technology.

Social informatics is "the interdisciplinary study of the design, uses and consequences of information technologies that takes into account their interaction with institutional and cultural contexts." (Kling, 1999). A key concept in social informatics research is that the *social context* of information technology

development and its consequences for work, organizations and other social relationships. For example, differences in expertise between users may radically influence their experiences with technology. Consider for example, the use of email filters to stop email from people known to be argumentative in a listserv community. A person who knows how to use these filters to stop the messages may develop a more favorable impression of the community than a person who receives the messages.

There is a strong need for people from different disciplines to collaborate in development of over-arching theories that will underpin usability and sociability.

Methods and measurement

Methods for designing online communities, investigating and measuring their effectiveness are starting to be discussed (Jones, 1999) (Preece, 2000). The size and demography of Internet communities is usually unknown making traditional research approaches, such as sampling, problematic.

Very few authors address community design issues; exceptions are Schuler and Figallo (Figallo, 1998; Schuler, 1996). An interactive, *community-centered* approach is needed (Preece, 2000). This method takes account of sociability as well as usability and emphasizes social community solutions as well as technical solutions. Knowing a community's purpose, the characteristics of participants and the nature of their communication tasks is an essential part of this approach.

For research and development (particularly collecting community requirements and evaluation) surveys, ethnography and data logging are promising methods. Survey techniques can be delivered through e-mail, web pages or by paper making them highly versatile and economical for researching geographically scattered populations (Lazar & Preece, 1999). Often combinations of all three approaches are needed to reach a representative sample of the population. As well as cost advantages data can be transferred into a database for immediate analysis and accessed via statistical programs such as SPSS, SAS or data visualization tools.

Ethnographic techniques have been well developed for CSCW research, but the distributed nature of many online community populations poses new challenges. One form of ethnographic research involves interaction logging, which is easy and unobtrusive (Nonnecke & Preece, 2000; Smith, 1999) but poses ethical questions for researchers. Effective ways of collecting and analyzing Internet data are also research issues. Algorithms are written to perform counts, for example, number of messages posted in a certain period of time, gender of posters (though assuming gender from login names is can be unreliable); maximum, average and minimum message length, thread length etc. This data can be treated statistically and also analyzed qualitatively to give a broader description of demography of online community life.

An important area of research both for demographic studies and to assess the efficacy of community design is to develop metrics to provide specific measures about demography of community members, their behavior and the characteristics of their interactions etc. Some typical metrics might be:

Demography of participants:
Number of members/subscribers in the community
Number of posters by gender, age, occupation (i.e. demography)

Behavior online:
Number of posts per person, connect hours etc.
Number of posters over a period of time (e.g. access to web pages per month)
Number of lurkers – but it is hard to get this information.

Characteristics of interaction:
Number of posters by category of type of communication
Length of message
Number of messages in a thread
Number of threads
How much and what type of moderation?

Summary and conclusions

It is essential that online communities research is multi-disciplinary and involves both social scientists and computer scientists. Software with good usability is important. But software alone will not ensure a successful online community. Sociability is also needed. This chapter has used these categories as a framework for reviewing research needs. In addition comprehensive theory and new methods for researching online communities are needed, which have also been discussed. A summary of the most significant suggestions follows which will form an agenda for future online communities research.

- *Usability*

Considerable research is needed to develop *representations* to reveal online behavior as it is occurring; histories of behavior, stored communication and knowledge, nature of communication (e.g. which topics were discussed), the number of people participating and relationships between participants.

Interfaces and interaction techniques are needed that are usable by a broad range of users for a wide variety of communication tasks.

Success of some online communities will be strongly influenced by how *secure* they are. Strong cryptographic protocols are essential.

Scalability is a research priority for online communities. With an increasing number of people from across the world developing and joining online communities, software and social processes are needed to support very large communities.

- *Sociability*

Research is needed by teams of social scientists to answer basic questions about social interaction and mass culture online. How do communities differ and what kind of software is needed to support them? What can we learn from physical communities that will enable us to develop better online communities? For example, how do daily, weekly and yearly cycles translate to online activity? What kinds of behavioral rituals and codes of conduct lead to successful online communities? Results of this research will lead to development of appropriate policies to guide online behavior and will also contribute to better usability design.

Ethical issues and universal access are significant research areas. For example, how to notify users visually about unencrypted or encrypted traffic, ensure privacy and encourage appropriate development of trust. Digital technology is impacting cultural diversity, environmental issues, conservation of limited ecological resources and people's standard of living. Research should inform national and international agencies so that they can deploy their resources well. Ways of providing universal access for people of all ages, cultures, languages, income levels, educational, physical and mental abilities.

- *Theory and methods*

Unifying theories influenced by sociology, psychology, social psychology, linguistics, communications research and psychotherapy are needed to inform online community development. A better understanding of sociability is needed urgently.

Researching, designing, measuring and evaluating online communities requires modification of well-established techniques and development of new one. Methods, techniques and tools are needed to measure online activities and to understand how online communities are different from geographical communities. A research agenda that encourages collection of demographic data enriched by ethnographic data will be particularly fruitful.

A good marriage of usability and sociability informed by strong theory and appropriate research and development methods will produce thriving online communities.

References

Ackerman, M., & Starr, B. Social activity indicators for groupware. *IEEE Computer, 29*(6), 37-42.

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

Brown, J., van Dam, A., Earnshaw, R., Encarnacao, J., Guedj, R., Preece, J., Shneiderman, B. & Vince, J. (1999b). Human-Centered Computing, Online Communities and Virtual Environments. *ACM SIGCHI Interactions, 6*(5).

Brown, J., van Dam, A., Earnshaw, R., Encarnacao, J., Guedj, R., Preece, J., Shneiderman, B. & Vince, J. (1999a). Human-centered computing, online communities and virtual environments. *IEEE Computer Graphics, 19*(6), 70-74.

Brown, J., van Dam, A., Earnshaw, R., Encarnacao, J., Guedj, R., Preece, J., Shneiderman, B. & Vince, J. (1999c). Special Report on Human-Centered Computing, Online Communities and Virtual Environments. *ACM SIGGRAPH Computer Graphics, 33*(3), 42-62.

Card, S. (1991). *Presentation on the Theories of HCI at the NSF Workshop on Human Computer Interaction* . Washington DC, USA: National Science Foundation.

Clark, H. H., & Brennan, S. E. (1993). Grounding in communication. In R. M. Baecker (Ed.), *Groupware and Computer-Supported Cooperative Work* (pp. 222-233). San Francisco, CA: Morgan Kaufmann Publishers.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*, 554-571.

Figallo, C. (1998). *Hosting Web Communities*. New York: John Wiley & Sons, Inc.

Friedman, B. (1997). *Human Values and the Design of Computer Technology*. Cambridge: Cambridge University Press.

Fukuyama, F. (1995). *Trust*. New York: Free Press Paperbacks, Simon and Schuster.

Gaines, B. R., Chen, L.-J. L., & Shaw, M. L. G. (1997). Modeling the Human Factors of Scholarly Communities Supported Through the Internet and the World Wide Web. *Jornal of the American Society of Information Science, 48*(11), 987-1003.

Garton, L., Haythornthwaite, C., & Wellman, B. (1999). Studying online social networks. In S. Jones (Ed.), *Doing Internet Research* (pp. 75-105). London: Sage Publications.

Hackos, J., & Redish, J. C. (1998). *User Analysis and Task Analysis for Interface Design*. New York: John Wiley & Sons.

Haythornthwaite, C. (1996). *Media use in support of communication networks in an academic research environment.* , University of Toronto, Toronto.

Jones, S. E. (1999). *Doing Internet Research. Critical issues and methods for examinig the net.* Thousand Oaks: Sage.

King, S. (1994). Analysis of Electronic Support Groups for Recovering Addicts. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century (IPCT), 2*(3), 47-56.

Kling, R. (1999). What is social informatics and why does it matter? *D-Lib Magazine, 5*(1).

Kollock, P. (1998,). Design principles in online communities. *PC Update, 15,* 58-60.

Lazar, J., & Preece, J. (1999). Designing and implementing web-based surveys. *Journal of Computer Information Systems, xxxix*(4), 63-67.

Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology, 63*, 234-246.

Liebowitz, J. E. (1999). *Knowledge Management Handbook*. Boca Raton: CRC Press.

Lynch, P. J., & Horton, S. (1999). *Web Style Guide (Preliminary Version)*. New Haven & London: Yale University Press.

Markus, M. L. (1987). Toward a critical mass theory of interactive media: Universal access, interdependence and diffusion. *Communication Research, 14*, 491-511.

Markus, M. L. (1990). Toward a critical mass theory of interactive media: Universal access, interdependence and diffusion. In J. F. C. S. (Eds.) (Ed.), *Organizations and communication technology* (pp. 194-218). Newbury Park: CA: Sage.

Morris, M., & Ogan, C. (1996). The internet as mass medium. *Journal of Communication, 46*(1).

National Telecommunications and Information Administration, U. S., & Commerce., D. o. (1999). *Falling Through the Net: Defining the Digital Divide.* . Washington, DC.

Nielsen, J., & Mack, R. L. (1994). *Usability Inspection Methods*. New York: John Wiley & Sons, Inc.

Nonnecke, B., & Preece, J. (2000). Lurker demographics: Counting the silent. *In press. Available from the authors.*

Norman, D. A. (1986). Cognitive engineering. In D. Norman & S. Draper (Eds.), *User-Centered Systems Design* . Hillsdale, NJ: Lawrence Erlbaum Associates.

Olson, G. M., & Olson, J. S. (1997). Research on Computer Supported Cooperative Work (Chapter 59). In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of Human Computer Interaction (2nd Edition)* (pp. 1433-1456). Amsterdam: Elsevier.

Preece, J. (1998). Empathic Communities: Reaching out across the Web. *Interactions Magazine, 2*(2), 32-43.

Preece, J. (1999a). Empathic Communities: Balancing Emotional and Factual Communication. *Interacting with Computers, 12*, 63-77.

Preece, J. (1999b). *A research agenda for online communities: Understanding behavior, predicting outcomes and informing design*. Bonas, France: EC and NSF Workshop.

Preece, J. (2000). *Thriving Online Communities: Usability and Sociability*. Chichester, England: John Wiley & Sons.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carrey, T. (1994). *Human-Computer Interaction*. Workingham: Addison-Wesley.

Rice, R. (1993). Media appropriateness. Using social presence theory to compare traditional and new organizational media. *Human Communication Research, 19*(4), 451-484.

Rice, R. (1994). Network analysis and computer-mediated communication systems. In S. W. J. Galaskiewkz (Ed.), *Advances in social network analysis* (pp. 167-203). Newbury Park: CA: Sage.

Rice, R. E. (1987). Computer mediated communication and organizational innovations. *Journal of Communication, 37*, 85-108.

Schuler, D. (1996). *New Community Networks: Wired for Change*. Reading, MA: ACM Press & Addison-Wesley Publishing Co.

Shneiderman, B. (1999). *Designing the User Interface: Strategies for Effective Human-Computer Interaction (Third Edition)*. Reading, M.A.: Addison-Wesley.

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. London: John Wiley and Sons.

Smith, M. A. (1999). Invisible crowds in cyberspace: mapping the social structure of the internet. In M. A. S. P. Kollock (Ed.), *Communities in Cyberspace* (pp. 195-219). London: Routledge.

Spool, J. M., Scanlon, T., Schroeder, W., Snyder, C., & DeAngelo, T. (1997). *Web Site Usability: A Designer's Guide*. North Andover: MA: User Interface Engineering.

Sproull, L., & Kiesler, S. (1991). *Connections: New Ways of Working in the Networked Organization*. Cambridge, MA: MIT Press.

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.

Wellman, B. (1992). Which types of ties and networks give what kinds of social support? *Advances in Group Processes, 9*, 207-235.

Wellman, B. (1997). An electronic group is virtually a social network. In S. Kiesler (Ed.), *Culture of the Internet* (pp. 179-205). Mahwah, N.J.: Lawrence Erlbaum Associates.

Whittaker, S., Terveen, L., Hill, W., & Cherny, L. (1998). *The dynamics of mass interaction.* Paper presented at the ACM CSCW'98 Proceedings, Seattle: Washington.