Lurking in email-based discussion lists

Robert Blair Nonnecke

A thesis submitted in partial fulfillment of the requirements of South Bank University for the degree of Doctor of Philosophy

March 2000

Acknowledgments

I would like to extend my gratitude to a number of people.

Any project of this magnitude requires a wide range of talents. I am indebted to Thawatchai Piyawat and Dick Seabrook for their consummate programming skills, and to Jantawan Noiwan for her statistical advice. Drafts of this thesis were considerably improved through the thoughtful comments and editing of Heather McDonald and Carolyn Davidson. Additional thanks go to mentors Tom Carey and Dave Hendry.

None of this would have been possible without the interviewees and the members of the discussion lists. I thank them for their time, insight and participation. The research was carried out at South Bank University and the University of Maryland Baltimore County. I am grateful to both universities for their support.

Finally, special thanks to my committee, Fintan Culwin, Xristine Faulkner, and Jenny Preece.

Abstract

The goal of this thesis is to increase our understanding of lurkers and lurking in online groups by addressing three *primary* questions: why do lurkers lurk, what do lurkers do, and how many lurkers are there? Lurkers reportedly make up the majority of members in online groups, yet little is known about them. Without insight into lurkers and lurking, our understanding of online groups is incomplete. Ignoring, dismissing, or misunderstanding lurking distorts knowledge of life online and may lead to inappropriate design of online environments.

To investigate lurking, the author carried out two studies. The first employed semi-structured interviews with members of online groups. This qualitative study addressed why lurkers lurk and what lurkers do. The second study, a log-based demographic study, examined the number of lurkers in discussion lists (DLs).

The ten DL members interviewed for the first study described 117 reasons for lurking, six major lurking activities and five key lurking strategies. It is clear that lurking is a strategic activity that involves more than just reading posts. Three models of lurking (*filter, gratification, and persistence*) were developed to account for lurkers' processes, needs, and circumstances. These models present lurking as an activity situated in the context of life both inside and outside of online groups.

The second study, carried out over a three month period, logged 147,946 messages from 60,000 members in 109 DLs. The percentage of lurkers was lower than expected (55% with no posts vs. 90% in the literature). However, when lurking was defined as three or fewer posts in three months, the level rose to 81%.

This thesis describes several other key findings. Health-support DLs were shown to have lower levels of lurking when compared to software-support DLs. The empathic nature of health-support groups may partially account for these lower levels. Another reason may be the lesser awareness among health-support members of the issues surrounding persistent messages, which may lead them to be less inhibited in their public posting.

Smaller DLs and DLs with shorter messages were found to have fewer lurkers. DLs with higher levels of interactivity were shown to have lower levels of lurking. Also, it was found that as traffic levels in DLs go up, lurking levels go down. This result flies in the face of the feedback from the interviews, which suggested that lurking is more likely to occur in high traffic lists. Obviously, something else is at work and several possibilities are suggested. Some lurkers experienced a sense of community while lurking. On closer inspection, lurkers meet many of the criteria of being community members, and this sense of community is not a surprising finding.

In order to clarify the term *lurker*, a new definition is needed The findings from this thesis are embodied in the following definition:

Extended definition of lurker: The term, *lurker*, is frequently used pejoratively and usually refers to anyone who never posts or posts infrequently. In fact, lurking is non-public participation. Lurking is a situated action, and many personal and group-, work-, and tool-related factors affect the activities and level of public and non-public participation. Lurking is "normal" in the sense that everyone is likely to be a lurker at some point in time. Lurkers are heterogeneous in most respects except in their lack of public posting. Therefore, in the absence of an understanding of the context in which it takes place, *lurker* is a meaningless term. Avoidance of the term *lurker* is recommended. Instead, the term *non-public participant (NPP)* is suggested. *NPP* is not pejorative and suggests there are other forms of valid participation outside of public posting.

Table of Contents

Page

Appendices

Index of Figures

Index of Tables

Chapter 1: Introduction

Overview

- The problem is explained.
- It describes the goals and the approach taken to address the problem.
- An introduction to lurkers in discussion lists (DLs) is provided.
- A description of how DLs function provides background information used throughout the thesis.
- The chapters ahead are detailed.

The chapter starts with a brief description of the problem and goals of this thesis. This is followed by an overview of what is known about lurkers in DLs. Following the overview is a section which explains how DLs function. This section has been included because understanding how DLs work is critical to understanding the thesis and its implications. At various points throughout the thesis, the functioning of DLs will be revisited to provide further clarification. The last section describes the chapters ahead.

1.1 Problem

Lurkers reportedly make up the majority of members in online groups and DLs in particular (Mason, 1999), yet little is known about who they are, why they lurk, what they do, or how widespread lurking is among online groups. Without examining lurkers and lurking, our understanding of online groups is incomplete. Ignoring, dismissing, or misunderstanding lurking distorts our knowledge of life online and may lead to inappropriate design of online environments.

1.2 Goals and approach

The goal of this thesis is to increase our understanding of lurkers and lurking by addressing three *primary* questions: why do lurkers lurk, what do lurkers do, and how many lurkers are there? To investigate these questions several methods are used. The first study employs semi-structured interviews with members of online groups. It addresses why lurkers lurk and what lurkers do. The second study. a demography, based on logged DL messages, examines how many lurkers there are in DLs. In the process of exploring these questions an improved and informed definition for the term, *lurker*, is developed.

1.3 Introduction to lurkers in DLs

Email-based DLs (aka listservs or lists), newsgroups, and Web-based bulletin board systems (BBSs) have experienced rapid growth as the number of Internet users climbs. As of July 1999, there are more than 131,000 DLs using Listserv's® server software. The 69,000,000 members of these DLs send in excess of 29,000,000 messages per day (L-Soft International, 1999b). Whittaker, Terveen, Hill, & Cherny, (1998) cite similarly large numbers for Usenet newsgroups. The growth and prevalence of online groups, coupled with the relative ease of gathering persistent and traceable messages, has made online groups a fertile ground for research. The following are a few of the areas so far studied: the development of friendship (Parks & Floyd, 1996), the perception and quality of community (Roberts, 1998), factors affecting interaction within newsgroups (Whittaker et al., 1998), and the development of empathy in health support groups (Preece, 1998; Preece & Ghozati, 1998). Each of these studies was based on examining individuals participating in public spaces, i.e., those who post. None examined their chosen area from a lurking

perspective, even though lurkers are reported to make up over 90% of online communities (Katz, 1998; Mason, 1999).

Given that lurkers are both unstudied and apparently in the majority, knowing more about them will have benefits in many areas. For example, their sheer number suggests they are an important area to study from an e-commerce perspective. As group development becomes an important component of commerce on the Internet, understanding lurkers will become an essential part of doing business. Many e-commerce enterprises present group facilities and a community oriented face, e.g., the Ask Dr. Weil Web site (Weil, 1999). Every lurker is a potential customer. For example, Amazon.com has been very successful in creating an online retail environment in which lurkers can make purchasing decisions based on how others have purchased in the past and on reviews supplied by other customers. Amazon.com has leveraged the information gained from those willing to post reviews into purchasing-support tools for the lurker and poster alike. From a usability perspective, improvements in tools and group design will fall out of a better understanding of lurkers and their activities. For lurkers and their communities, knowledge of lurking will have the benefit of demystify lurkers' roles, value, and activities. This has already shown to be the case. The following is how one lurker responded to an initial draft of an article on lurking (Nonnecke & Preece, 1999):

> Maybe it's a sign of my own mild discomfort around being a lurker, but I found it reassuring to recognize myself and my behaviour within the continuum you describe, and to see lurking treated seriously, with both acceptance and respect. As a lurker, I'm used to observing from the sidelines and participating vicariously, and it's strangely gratifying to read an article that speaks directly to that experience. It's almost like suddenly feeling part of an (until-now) invisible community of lurkers.

Researchers also have opinions about lurkers that need to be verified. For example, Kollock and Smith (1996) describe lurkers as "free-riders", i.e., noncontributing, resource-taking members. Knowing more about lurkers and their lurking will show whether this is an accurate description.

Definitions for *lurker* and *lurk* provide insight into how lurking is viewed. The online Jargon Dictionary (1999) defines the term, *lurker*, as:

> One of the 'silent majority' in a electronic forum; one who posts occasionally or not at all but is known to read the group's postings regularly. This term is not pejorative and indeed is casually used reflexively: "Oh, I'm just lurking." When a lurker speaks up for the first time, this is called 'delurking'.

This definition suggests that lurking is the normal behaviour of the majority of the population and that lurking can be defined in terms of the level of participation, either as no posting at all or as some minimal level of posting. In contrast to the Jargon Dictionary, Merriam-Webster's WWWebster Dictionary (Merriam-Webster, 1999) provides a pejorative definition for the term, *lurk*:

> **a :** to lie in wait in a place of concealment especially for an evil purpose **b :** to move furtively or inconspicuously **c :** to persist in staying

These contrasting perspectives reflect an inadequate understanding of the lurker in online discussion forums. The former definition evokes the image of a benevolent yet responsible Net citizen, while the traditional definition implies something much more sinister. Evidence for the former is anecdotal and, without appreciating the nature of online lurking, the latter definition may be inappropriate.

Defining lurking is problematic. Should someone who never posts in public spaces but regularly side-posts to individual group members be deemed a lurker? If a person posts once and then never again, does that constitute lurking? Is someone lurking when they go on holidays? Is someone lurking when for a period of time they do not post? While these are important considerations, this thesis takes an initial, simple approach of defining lurking as either no posts or some minimal number of posts over a period of time. A goal of this thesis is to systematically improve this definition using empirical evidence. The improved definition will

have the practical value of publicizing an improved understanding of lurkers and ensuring that electronic discussion groups effectively serve lurkers.

Lurkers undoubtedly exist in many online environments. However, for the purpose of this thesis, lurkers in DLs will be the focus. DLs were chosen as the group environment because of their popularity. Just as importantly, membership levels can be determined in DLs, something which is difficult in other asynchronous group environments such as newsgroups and bulletin board systems (BBSs). Knowing membership levels is crucial to determining lurking levels as the number of lurkers in a DL is the total membership minus the number who post. In the case of DLs, those who post can be counted by tracking the authors of the posted messages. In order to understand the remainder of this thesis an understanding is required of how email-based DLs work. The next section provides an overview of how DLs function.

1.4 How DLs function

DLs are automatic devices for sending and receiving messages amongst members of a group. They are also asynchronous communication tools in which members can choose when to view their messages, if at all. DLs facilitate delivery of email to a set of subscribed members using a broadcast model. Anyone who sends email to the central server effectively broadcasts the email to all members of the DL. Individuals can respond to received email via the server, which in turn broadcasts the reply to all members. There may be an intermediate step in which messages are moderated. This can introduce delays in propagation and/or the elimination of some email, depending on how the moderation is handled. At the member's option, the email may be received individually or in the form of a digest (a group of messages).

An important aspect of DL messages is their persistence. In this case, "persistence" means the continued availability of messages, often for an indefinite period of time, and not only in each member's email storage but also in private, public, and corporate locations. For example, many DLs keep all messages in a central public archive that can be easily searched. Email may also be intercepted or backed up and held in corporate databases. Being both persistent and dispersed means DL messages are searchable and manipulable, and available to non-group members.

DL messages contain header information that includes sender, date, and subject. The header information allows messages to be sorted and managed using a variety of software. (For example, users of the email client Eudora can follow a thread by sorting messages by author, subject, and date.) In addition, each message contains content and, frequently, a signature. Both the header and message content make great fodder for searching. Searching can range from members searching their own locally maintained email to a researcher searching for quoted text through the use of crawler-based search engines, e.g., Excite (Excite.com, 1999). Because copies of messages may reside in many locations outside of the subscriber's control, access is effectively wide open. The messages can be searched for content, originator, or in many other ways.

DL email may be read in isolation and the flow and intent of the messages can be distorted through the redistribution of individual messages or parts of copied messages. For all intents and purposes, email from DLs may be mutated from dialogue to data and back to content, without the originator having control over the process or use. An example of unintended use is the trolling of DLs for the purpose of creating address lists, which are then sold to spammers and legitimate businesses. The copies may also be used in the way they were intended, e.g., as an accessible resource for the group, for finding specific information, and for following conversations.

At the individual level, people manage their incoming email, including their DL email, in many different ways (Whittaker & Sidner, 1996). Some people have high volumes of email while others have low volumes. Some people read all messages and others do not. Some people file messages while others keep all their messages in a central inbox. How people manage their email has an effect on how they manage their DL messages. As a result, it is unlikely that all members of a DL see or treat their messages in a uniform manner.

In addition to the DLs, group members are frequently supported through related Web sites, sets of frequently asked questions (FAQs), and electronic forums such as chat rooms and bulletin board systems (BBSs). Web sites are becoming a common gateway to a number of different information and communication tools, e.g., chat spaces and BBSs.

1.5 Chapters ahead

The *problem* is that little is known about lurkers in online groups. As lurkers reportedly constitute a majority of the online population, knowing more about them is important to the general understanding and design of online groups. The *goal* of this thesis is to shed some light on lurkers and their activities in DLs.

Three *primary* questions (Ps) to be developed in the next chapter are essential to the understanding of lurkers and lurking.

P1: Why do lurkers lurk?
P2: What do lurkers do?
P3: How many lurkers are there?

These questions will be used to guide much of the remainder of the thesis. In addition to the *primary* questions, a set of *related* questions is developed in reviewing the literature on lurkers and online groups in Chapter 2. The *primary* questions are then used in Chapter 3 in a review and assessment of six different research methods. Chapter 4 contains a description of and results from the first of two studies. The first study uses semi-structured interviews to address P1 and P2. A discussion of the results is presented in Chapter 5. Chapter 6 presents a demographic study in which messages are logged from 109 DLs. This study focuses on P3 and contains both the results and a discussion of the results. Results from both studies are used in Chapter 7 to develop a coherent picture of lurkers in DLs. Chapter 8 contains the conclusions and recommendations for future work.

Chapter 2: Literature Review

Overview

- The first section reviews the literature for references to lurkers in online groups and finds a paucity of information.
- A review of two primary articles on lurking raises both *primary* and *related* questions.
- Five of the *related* questions are used to organize a review of the pertinent literature about online groups.
- The importance of studying online groups and DLs is established.
- The chapter is summarized and the direction forward is outlined.

The first chapter described the "lurker" problem and the goal of this thesis. This was followed by an introduction into lurking in DLs and a description of how DLs function. The purpose of this chapter is to review the literature related to lurking and online groups and to establish a series of questions that will be addressed in the remainder of this thesis. The *primary* questions raised in this chapter are used in the next chapter to evaluate methods for researching the "lurker" problem.

Very little work has been published on lurkers in online groups. Therefore, to give a sense of what is being said about lurkers, the first section of this chapter is a sampling of comments on lurkers. The general finding from this first section is that not much is known about lurkers and they are worth studying. The second section reviews two articles in which lurkers are discussed directly and in very different ways. The first article is from an online publication and the second is a conference paper. In the course of reviewing the literature in the first and second sections, a series of *primary* and *related* research questions (Ps and Rs) are put forward. The third section in this chapter uses the *related* questions to organize a review of the literature on online groups as it relates to lurkers. Answers to the *primary* and *related* questions will also be sought in the two studies described in subsequent chapters (Chapters 4, 5, and 6). The fourth section of this chapter examines the importance of studying online groups and DLs.

There are numerous questions related to lurking. However, though many will be brought up in this and later chapters, it would be impossible to address all of them in this document. As a preview of the questions that will arise in this chapter, the following lists the *primary* and *related* questions that will be used to focus this inquiry:

P1: Why do lurkers lurk?
- R1a: What motivates lurkers?
- R1b: What role does lurking play in learning about the group?
- R1c: How does persistent conversation affect lurking?
- R1d: How do individual and group character differences affect lurking?

P2: What do lurkers do?
- R2a: What are the constraints on lurkers' activities?

P3: How many lurkers are there?

Note: These questions will not be raised in the above order , but will be elicited based on the literature reviewed.

2.1 References to lurkers in online groups

From reviewing the literature, it is apparent that lurking and lurkers in online groups have not been studied extensively. This section describes references, beginning with a short rationale for why lurkers have not been studied and why they should be.

Researchers have speculated on the nature of lurking. In his work on networked interactivity in online groups (Rafaeli & Sudweeks, 1997), Rafaeli defines interactivity as the "dependency among messages in threads". The underlying assumption is that the measure of interactivity is based on the examination of the public messages. As lurkers by definition post infrequently or not at all, this is a measure that excludes lurkers and their contribution.

This exclusion of lurkers is recognized by Rafaeli. Without claiming an importance for, or even a need to study and understand lurkers and lurking, he makes the following comments: "There is a silent portion of participants about whom we can only speculate." and further on "We have no information about unverbalized reactions (of the lurkers)…". Rafaeli's first statement describes lurkers as participants, and his second indicates that little is known about them. Describing lurkers as participants, silent though they may be, suggests that they participate in unknown ways. This in turn suggests that knowing how and why they participate is an important aspect of studying online groups.

The lack of current information about lurkers is due in no small part to the methods used to study online communities. These methods frequently rely on public messages as the primary raw data. In their review of the Internet as a form of mass media, Morris and Ogan (1996) point out the paucity of information on lurkers. They ask the following questions about lurkers, their number, and their nature:

> We may discover a fair amount about the producers of messages from the content of their electronic messages, but what about the lurkers? Who are they and how big is this group? To what extent do lurkers resemble the more passive audience of television sitcoms? And why do they remain lurkers and not also become information providers? Is there something about the nature of the medium that prevents their participation?

In their extensive log-based study of mass interaction in newsgroups, Whittaker et al. (1998, p. 263), indicate that lurkers need to be studied. They would like to know the rationale for lurking, whether lurking is transitory, and whether group dynamics are a factor:

> Why do people contribute to certain discussions but not to others? How long do people lurk before they first post? And how is dominance viewed? Why do certain people post multiple messages and how are they perceived by others for doing so?

The method employed in their study consisted of counting the observable, i.e., counting public posts in newsgroups over a six-month period. Because they used newsgroups, whose readership is impossible to determine, as the basis of their study, they were unable to determine the number of lurkers Like Rafaeli, Whittaker et al. recognize that lurkers are part of the equation, even though the methodology employed precludes their study.

Individual researchers have characterized lurkers in a number of ways. For example, lurkers have been described as communicationally incompetent, i.e., "people who lurk do so because they do not feel competent to post" (Mason, 1999). This observation comes out of Mason's ethnographic study of British football fans. It represents his perception of lurkers in a specific community and may or may not apply to different types of groups. It is also an observation that does not appear to be based on discussions with a wide range of lurkers within that community. Part of the problem he found in coming to understand lurking in this DL was that most lurkers are by nature less open to being studied. In addition, he employed what he calls "virtual ethnography", i.e., his ethnographic study took place through the Internet. This in itself may impose a different set of barriers to ethnography than face-to-face ethnography. One such barrier might be an increased difficulty in luring the lurker into participating in the ethnography. By contrast, lurkers are more than willing to describe their lurking in face-to-face interviews, as will be seen in the first study of this thesis (Chapter 4).

In addition to their lack of confidence in their competence, lurkers may exhibit the kind of passivity commonly associated with TV viewers ((Morris & Ogan, 1996) and Postmes, personal communication, 1998). They have also been characterized as abusers of the common good, i.e., [lurkers do] "…not contribute to the joint effort, but free-ride on the efforts of others" (Kollock & Smith, 1996). Lurkers as free-riders will be examined in more detail in the next section. How lurkers are viewed versus how they view themselves raises a *related* question connected to what lurkers derive from lurking:

R: What motivates lurkers?

What evidence there is, for the level of lurkers in online groups, suggests they make up a sizable majority. Reports of lurking levels range from 90% (Mason, 1999) to 98% (Katz, 1998). In Mason's ethnographic study, the reported level is for one specific community, while Katz reports an aggregate value for a number of DLs in which the author was involved (more on this in the next section). While the numbers are high, the variables affecting the lurking levels, or whether groups differ in their overall lurking levels, is not known. Even at this early stage of the literature review, it is apparent that knowing how many people lurk would be useful information. The numbers provided by Mason and Katz suggest very high levels of lurking. It is important to verify these numbers as they suggest that lurkers are an extremely important part of online groups, if only because they represent a large majority of the membership. As a result, one of the *primary* questions for this research is:

P: How many lurkers are there?

It is apparent that researchers do not understand lurkers and their activities. That coupled with lurkers apparent majority in online groups strongly suggests that not only are they important to understand, but that the current research orientation and methods have failed to illuminate the lurker. Methodological issues will be addressed in the next chapter, where orientation and methodologies currently used to study online groups will be discussed.

While very little has been published on lurkers, two articles address lurking directly. One is a column describing one person's intimate experience with lurkers and the other is a description of how lurkers can be viewed from a theoretical perspective as "free-riders". Both are reviewed in the next section.

2.2 Review of two primary articles

The first of two articles to be reviewed in this section is by Katz, a feature contributor for the Web site Slashdot: News for Nerds. Stuff that Matters (Slashdot, 1999). In a column titled Luring the Lurkers (Katz, 1998), Katz describes his understanding of lurkers based on both the email he receives after each column - between 100 and 500 emails, much of it from lurkers - and his observations of online forums. Katz describes lurkers from three perspectives: who they are; their reasons for lurking; and their value.

After a column describing his difficulties with learning Linux, email from lurkers was overwhelmingly supportive. At the same time he was receiving this support, messages in a related public BBS were highly critical. He indicates that the email he receives from lurkers can be challenging, but is not hostile in the same way it is in BBSs. His lurkers are tolerant to open discussion and are technically sophisticated. They also come in a wide variety of types: young, old, men, women, interested, disinterested, etc.

Katz has come to understand why his lurkers lurk: they are uncomfortable with the tone and hostility of public forums, and they believe that the values espoused in the public forum are widely held and they are alone in their opinions. Lurkers who would like to post desired moderated discussions which ban anonymous posting and personal insults. As well, non-native English writers lurked out of a lack of confidence in their English skills This multitude of reasons for lurking is in contrast to the single reason (feelings of incompetence) stated by Mason (1999). The single reason supplied in the research literature suggests that this is an area not well understood. This leads to another *primary* question:

P: Why do lurkers lurk?

About one-third of Katz's lurkers prefer lurking, "…bypassing the worst, personal insults, and abuse…" Of those who did not prefer lurking and were interested in posting, reasons of a social nature were cited for their lurking, e.g., to become familiar with the terminology and rituals used in the public forums. While Katz talks about their technological sophistication, it is not clear whether this is also a reason for their lurking, e.g., fear of being tracked through persistent conversation.

The overall message of Katz's article is that lurkers are to be valued and not shunned. Lurkers are rational, less bellicose participants who lurk for a variety of reasons. Katz concludes with a call to lure out lurkers by creating WWW sites that address the needs of lurkers and welcomes their input. The underlying belief is that lurkers are valuable to the community and that online groups could be a better place by making these communities more inclusive. Katz's work may be limited by the focus of his particular community – a group interested in things technical, e.g., learning about installing a UNIX operating system - and as such, may not reflect the dynamics of groups with a different focus, e.g., health support. While Katz calls for luring the lurker, how this might take place is still up in the air. In order to understand how this might be done a *related* question needs to be answered:

R: How do individual and group differences affect lurking?

Most studies treat lurkers as if they are a homogeneous group. In the second article to be discussed in this section, Kollock and Smith (1996) describe online lurkers as free-riders. Their work is built on Ostrom's earlier work on face-to-face communities (Ostrom, 1990). Kollock and Smith provide a top-down framework as applied to Usenet newsgroups. While their framework is steeped in the tradition of face-to-face relationships, no empirical evidence is provided to support their conclusions or framework. Nor is it clear why Ostrom's earlier framework describing face-to-face relationships should apply to online groups.

Kollock and Smith give the following description of the "free-rider problem":

> Whenever one person cannot be excluded from the benefits that others provide, each person is motivated not to contribute to the joint effort, but to free-ride on the efforts of others. If all participants choose to free-ride, the collective benefit will not be produced. The temptation to free-ride, however, may dominate the decision process and thus all will end up where no one wanted to be.

Included under the free-rider umbrella are lurkers and abusers of decorum or bandwidth, e.g., people who flame, cross-post to other discussion forums; or post off topic or post large messages. Inclusion of lurkers and abusers under the same category makes some of the arguments difficult to follow as many of the points they make for decorum can be achieved through lurking, e.g., understanding rules before publicly posting. The rationale for lurking is not addressed, merely the effect it has within the socio-economic model. Nor is a range of lurking activities addressed, e.g., lurking as a preferred way of participating, and lurking as a means of learning the terminology and rules (as described by Katz). It is unclear whether the authors have studied lurkers as no studies were reported. Without empirical studies, this is a somewhat speculative framework.

In describing their socio-economic model Kollock and Smith state that "the more people free-ride, the more difficult it is to produce useful information and interaction." The authors appear to be suggesting that the focus of an online group is for the production of "useful" artifacts rather than as a means of communication as argued by Kraut et al. (1998b). The term "useful" is problematic, especially as a means of measurement, as useful can mean many different things depending on whose interests are at stake. Whether lurkers view DLs as producing useful artifacts or places for communication is unknown. This is an important distinction as each could carry a different set of responsibilities. For example, lurkers could be a successful audience and thus fulfill their communication obligation, but if the emphasis is on producing artifacts, then the lurker may have difficulty contributing to its public production. If we abide by the framework, then non-public artifacts, such as side posts, may play a role for lurkers in fulfilling their duties in producing useful artifacts. The lack of understanding of what lurkers do, leads directly to another *primary* question:

P: What do lurkers do?

Without this understanding, it is difficult to conclude whether lurkers are free riders.

With regard to the personal costs of making public posts, Kollock and Smith state the following: "communicating with thousands of people has essentially the same personal cost as sending a message to a single individual". Personal cost in this context appears to be an economic one, and does not take into account the anxiety, at least for some people, associated with voicing an opinion or making a statement in public. "Fear of public speaking consistently tops every list of human fears" (Wilder, 1999). It would be no surprise if a similar fear exists in online environments. If the personal costs of posting in public and private are the same, side posts should be roughly equivalent in number to public posts. That is, to be consistent with their activity in public spaces, lurkers would not side post, and regular public posters would also side post. Based on Katz observations, neither appear to be the case. Katz's email contained many more lurkers than non-lurkers. For some of Katz's lurkers, there is apprehension over public posting.

In discussing their work as related to newsgroups, Kollock and Smith bring up an important difference between how newsgroups or Web-based bulletin board systems (BBSs) operate and how a typical email-based DL works. In describing messaging in a newsgroup, they make the following statement: "a great number of members can participate in discussions involving numerous topics without overloading participants". In both newsgroups and BBSs, the group member goes to the group either through a news reader or the BBS user interface (UI), and selects messages which are separated into threaded dialogue. In this manner, the member can view messages for a particular group, isolated from other groups, and at their leisure. In DLs, because messages are often received in a common inbox with other email, including messages from other groups, there is a potential danger of overloading the member. The danger is many fold; the messages from different groups will compete for attention, and the visual and software distinctions afforded in both newsgroups and BBS is not automatically available in an email client. This suggests that there may be a danger in overloading DL members under certain circumstances, e.g., high volume of email and long conversations where headings change. This brings up another *related* question:

R: What are the constraints on lurkers' activities?

Kollock and Smith describe persistence of conversation in newsgroups as a positive attribute. However, it has been noted by others, e.g., Erickson (1999) that persistent conversation is a two-edged sword. It provides members with the ability to find information, but it also provides the means for others to search the information and use it however they wish. Whether members know this and whether this is an impediment to posting is not known. Persistence is such an important part of online interaction that understanding persistence as it relates to lurking leads to the *related* question:

R: How does persistent conversation affect lurking?

The relevance of persistent conversation to lurking is discussed in more detail in Chapter 5, where it will be used to understand the results from the first study (Chapter 4).

Kollock and Smith also describe the value of ongoing interactions in a group: "knowing that one will be interacting with others on a continual basis can lead to the creation of reputations and serve as a powerful deterrent to short-run, selfish behavior". One of the selfish behaviours the authors refer to is lurking. Without empirical evidence, it is unclear whether reputations work in this manner, especially in the case of lurkers. For example, Katz suggests that reputations built on pedantic or aggressive rhetoric have the effect of keeping lurkers lurking. Empirical evidence showing the distribution of posters and the number of posts within DLs would be a useful tool in understanding the role of reputation with regard to lurking. If the distribution of posts among posters showed that a few posters were dominating the conversation, it might be expected that lurking would be higher.

Katz's lurkers talked about coming to understand the social norms (rituals) of a group. In a similar vein, Kollock and Smith talk about the rules of a group:

> Any successful community will have a set of rules -- whether they are implicit or explicit -- that govern how common resources should be used and who is responsible for producing and maintaining collective goods. However, it is important that the rules are tailored to the

> specific needs and circumstances of the group. Ostrom identifies this as another design principle that is a feature of cooperative communities: there is a good match between the goals and local conditions of a group and the rules that govern the actions of the group's members. Her research indicates that there is often great variation from community to community in the details of the rules for managing collective goods. One lesson is that it is dangerous to take the specific rules of a successful group and apply them blindly to other groups.

Rules are important, whether they are set down in writing or are embedded in the interaction between members. Differences among communities may show up in their various rules of interaction, and that for the collective good, members need to understand these rules. There may be a similarity between how people learn to use software and how they come to learn the rules of a community. It has been well documented that many users of software do not read the instructions. In a similar manner, DL members may not read the DL's rules which accompany new subscription notices. It is unlikely that a set of explicit rules for a group can be all-encompassing especially when many of the rules may be tacit, and only understandable in the context of specific interaction between members. Lurking could be the equivalent of looking over someone's shoulder as they demonstrate how to use a piece of software. In this light, Katz's lurkers appear to be good citizens rather than free-riders. Knowing more about how lurkers use their lurking to understand communities leads to the next *related* question:

R: What role does lurking play in learning about the group?

Both of the articles reviewed in this section point out many possible reasons for lurking and that understanding why people lurk in online groups will contribute to both an understanding of lurking and to an understanding of online groups.

Kollock and Smith's article has influenced other researchers, many of whom have used the term, *free-rider*. For example, Wellman and Guila (1999) in their discussion on whether virtual communities are communities, make reference to Katz and Kollock's work. Wellman and Guila propose that "free-riders" lurking in support groups are less detrimental than in face-to-face situations because their lurking is not as easily observed. In their discussion of BBSs, Morris and Ogan (1996) talk about a "critical mass" of users required to carry the "free riders". They go on to talk about "members, participants, or free riders" in a way that suggests that participation is strictly defined as posting in public spaces. No mention is made of participation in other ways, such as direct email between members or other forms of communication or relationships.

In summary, the two articles (Katz, and Kollock and Smith) provide contrasting opinions from very different perspectives. On one hand, Katz views lurkers as participants who should be encouraged to participate within communities. His experience in online groups and the email he has received from lurkers is the basis for this position. Katz's work is almost ethnographic in quality. On the other hand, Kollock and Smith indicate that lurkers free-ride on the efforts of others. They appear to have concluded this by adopting and extending another researchers framework and do not support these findings with research.

All in all, surprisingly little is known about lurkers in online groups. Given lurkers apparent large numbers, this is something of a mystery, though it may be possible to explain the lack of research based on a number of issues. These include the difficulty of studying lurkers; the fact that this is a relatively new field and other hot topics have been bigger research draws; and that lurkers may not have received much attention because they lurk and therefore, their presence is easy to overlook.

Three *primary* and five *related* questions were developed in this and the previous section:

P1: Why do lurkers lurk?
    R1a: What motivates lurkers?
    R1b: What role does lurking play in learning about the group?
    R1c: How does persistent conversation affect lurking?
    R1d: How do individual and group character differences affect lurking?
P2: What do lurkers do?
    R2a: What are the constraints on lurkers' activities?
P3: How many lurkers are there?

Given the paucity of direct research on lurkers, the next section will examine research into online groups, with an eye to understanding what it might have to say about lurking. The five *related* questions are used to organize the literature review presented in the next section.

2.3 Literature organized around five *related* questions

In the previous sections, a number of articles discussing aspects of lurking were reviewed. From those reviews, the *primary* and *related* questions were put forward. This section contains a review of the literature as it pertains to the *related* questions.

R1a: What motivates lurkers?
R1b: What role does lurking play in learning about the group?
R1c: How does persistent conversation affect lurking?
R1d: How do individual and group character differences affect lurking?
R2a: What are the constraints on lurkers' activities?

R1a: What motivates lurkers? In a paper/discussion on why communication researchers should study the Internet (Newhagen & Rafaeli, 1996), Rafaeli suggests that gratification is an important element in understanding why people put considerable time and effort to connect over the Internet. He questions why people expend so much effort presenting themselves and then suggests that interaction between members is likely to play a major role. Trying to understand lurking in this context is confounding. Lurkers do not publicly present themselves, and public interaction for the lurker is unidirectional with only half of the gratification possible, that of being a recipient. The fact that online group members lurk, suggests that connecting may not be the sole source of gratification or even the most important. If Rafaeli's suggestion is true, that gratification is a strong motivation, then lurkers will likely have sources of gratification outside of the direct connection.

In their discussion of the Internet as mass medium, Morris and Ogan (1996) talk about receivers, or audience, for messages, and that these receivers "may or may not move fluidly from their role as audience members to producers of messages". The use of the term "role" and "audience" suggests a passivity on the part of group discussion participants. However, no evidence is cited to support this conclusion. Their description of fluid movement from audience to producer implies there are reasons/motivations for the change. As well, they do not indicate whether this movement is bi-directional, or whether it is only from audience to producer.

R1b: What role does lurking play in learning about the group? In their study of mass interaction in newsgroups, Whittaker et al. (1998) suggests that the activities of lurkers are a legitimate form of participation, i.e., a background involvement that can be beneficial. They support this position by citing Kraut and others, who see this as an important transition mechanism for novices to learn about a novel topic (or social milieu). In describing members of social groups, Gunnarsson (1997, p. 148) indicates that the members "are shaped or socialized with respect to knowledge, norms, attitude, and identity". It is likely that at least some portion of lurking behaviour is attributable to the process of coming up to speed on the workings of a group. This process may require more observation and listening and less public participation.

Beaudouin and Vekovska (1999) describe the building of identity and the taking on of roles and status within a newsgroup-based community. In their study of the Cyberian newsgroup they provide an ethnographic account of how relationships were built. They found that regulars in the group had a sense of

belonging when talking about newcomers who did not catch their jokes. This exclusiveness and bonding was undoubtedly recognized by those new to the group. Although their study is not directed at lurkers per se, it would not be surprising to find that lurkers recognized that they were outside of the core group. In this situation, lurking would be an obvious way of learning about the group without putting oneself at risk. The authors found that community members value one-to-one relationships and they used many other channels outside of the newsgroup for communication, e.g., email and ICQ (an online communication tool that combines both asynchronous and synchronous communication capabilities).

Related to the larger area of social cues is online identity and the construction of self through the use of signature, Web sites, and other means. The paucity of online social cues like body language or gender has been held up as a barrier to communication in online groups, and a reason to be dismissive about online communication. However, studies have shown that reduced-cues are not as important as they were thought to be. In their study of 100 online support groups, Preece and Ghozati (1998) report that empathy is prevalent and well communicated in many online self-help health groups. It is now generally agreed that reduced cueing may slow the speed of developing relationships rather than being a complete barrier. In her seminal work on the design of online social environments and communities, Donath (1996) suggests that readers of newsgroups seek the identity of those giving advice, and that this is done in several ways, i.e., through reputation, signatures, and archives. The literature does not address DLs specifically, but there is nothing to suggest that DLs operate any differently.

A number of researchers have been working on showing social activity and presence (Ackerman & Starr, 1995; Ackerman & Starr, 1996; Donath, Karahalios, & Viegas, 1999; Viegas & Donath, 1999). In DLs, presence can be determined by examining postings or by querying the DL server for a list of members (this is becoming less common as the default on new DLs is to disallow this type of query). Those who belong to the DL and want to remain anonymous and unidentifiable can lurk. Presence of the group as a whole is not displayed in DLs, e.g., the number of people reading a message is unknowable. Work on showing both presence and activity in synchronous environments has been undertaken by Viegas and Donath (1999). However, this work has yet to be applied to primarily asynchronous environments or where the UI is variable and diverse, i.e., in heterogeneous email clients. Ackerman and Starr (1996) argue for social indicators, explaining that “people pay a great deal of attention to the activities of others”. They argue that the number of social indicators can be extremely varied, not only at the interface level, but also based on the different needs of individuals. One of the arguments they make for social indicators is that members will more closely attend to systems when they are aware of interesting activity. In a crude sense, DL email dropping into one’s mail box is a social activity indicator.

Parks and Floyd (1995) examined the development of friendship in newsgroups. They polled a large number of regular newsgroup participants to determine whether participants developed friendships. They found that friendship is possible and a frequent feature of public membership in newsgroups. This study was based on examining those who post, so it is unclear whether their findings would extend to lurkers. They found that when friendships developed, they involved contact outside of the newsgroups in the form of email (98%), telephone calls (35%), face-to-face meetings (33%) and correspondence sent via the postal service (28%). These results and those described in the previous study of Cyberians suggest that observable public participation is the proverbial tip of the iceberg and that non-public behaviours may account for a large portion of the group’s interaction.

R1c. How does persistent conversation affect lurking? Based on Erickson’s definition and description (Erickson, 1999) email is a persistent medium. The following description is from his call for participation in the Persistent Conversation mini-track at the Thirty-Third Hawaii International Conference on System Science:

> [persistent conversations] include conversations carried out using email, mailing lists, news groups, bulletin board systems, textual and graphic MUDs, chat clients, structured conversation systems, document annotation systems, etc. The persistence of such conversations as computerized records, although variable in duration and ease of user access, gives them the potential to be searched, browsed, replayed, annotated, visualized,

> restructured, and recontextualized, thus opening the door to a variety of new uses and practices.

Whether users of the technology grasp how persistent conversation can affect them is unclear. Lurkers may or may not be aware of persistence as an issue, but are likely to encounter it in various ways. For example, persistence may create apprehension over how a public dialogue could be taken out of context at a future date, and at the same time provides a means of asynchronously following threaded conversation. Persistence also allows conversations to be followed long after they have taken place through the use of archives. Donath et al. (Donath, Karahalios, & Viegas, 1999) have been working on ways of visually presenting persistent threaded conversation. One of their goals is to help readers "comprehend the discussion's structure and history and become familiar with its community". The underlying notion is that becoming familiar with a group is an important activity and that persistent conversation is a substrate for providing that familiarity. In the case of DLs, messages in archives and in email clients are the persistent messages.

R1d: How do individual and group character differences affect lurking? At the individual level, gender differences have been found to be a predictor for making friendships in newsgroups (Parks & Floyd, 1996). Women made friendships significantly more often, but the reasons for this difference are not known. (On the other hand, age has not been found to be a predictor for making friendships.) In her work on whether newsgroups are virtual communities, Roberts examined gender issues (Roberts, 1998). She found that women had a greater sense of community. Women also created the majority of the posts (75%) in newsgroups with high levels of posting (>80/day). Preece and Ghozati (1998) found similar results in their study of empathy in health-support groups.

One aspect of group character is how a group communicates. For example, Beaudouin and Velkovska (1999) found that exclusion from a group can be done by non-response. The exclusion can take the form of excluding certain topics or can be directed at specific members. Beaudouin and Vekovska call this "symbolic violence", and it may be noted and perhaps feared by lurkers. A less symbolic form of violence is flaming, which has been shown to be fear-inducing (Kayany, 1998). Kayany investigated the social context of flaming in newsgroups and found that the newsgroups of different topics have different rates of flaming. Knowing whether lurkers fear flaming or symbolic violence may be an important step in understanding why lurking occurs. Similarly, knowing whether members are drawn to lurk or delurk in empathetic groups such as those described by Preece and Ghozati (1998) is not known.

Dialogue has been used to determine the degree of interaction in online groups. A six-month logging study of several hundred newsgroups (Whittaker et al., 1998) showed that short messages correspond with high levels of threading, i.e., greater interactivity. This may be a result of how news readers function (or for that matter how messages are read online), i.e., short messages are easier to read online. Other factors may be at work, such as the difficulty in keeping a thread focused when long messages are involved. Threading is poorly shown in email clients and some DLs are distributed in digest form, making it much more difficult to follow threads. As a result, shorter threads may occur in DLs.

A major influence on group character comes from the moderation applied to a group. Moderation can come in many forms, from censoring of messages and removal of members, to a gentle nudge when a topic heading change is called for (Berge, 1992; Collins & Berge, 1997). In reaction to issues related to censorship and academic freedom in scholarly discussion groups, Berge developed a list of moderator roles:

- Facilitator (keeps list "on track"; acts as group leader)
- Manager (acts as administrator, archives messages, adds and deletes subscribers)
- Filter (decides upon on-topic posts; increases signal:noise ratio; deletes libelous posts; may delete jokes)
- Expert (answers frequently asked questions, acts as expert in the list's field)
- Promoter (asks questions of the list subscribers to promote discussion)
- Marketer (promotes/explains list to potential subscribers)
- Helper (helps people with needs – more general than expert)
- Fireman (takes "flames" or ad hominem attacks offline)

The impact of the various forms of moderation on participation is largely unknown.

R2a: What are the constraints on lurkers' activities? There are a number of constraints on group members that may affect their participation and thus their lurking. For example, the amount of time available for participating in online groups will vary from member to member. The following was noted in Parks and Floyd's work on developing online friendships: "Walther and his colleagues found that the proportion of socioemotional content was higher when interaction time was not restricted." (attributed to Walther, Burgoon, & Park, 1994). If the result of lurking is thought of as a lowering of the visible socioemotional content of a group, then it may be because lurkers have less time available to publicly participate.

Communication overload in email clients has been studied extensively (Whittaker & Sidner, 1996; Whittaker et al., 1998). One suggestion is that long messages cause communication overload, and that short messages promote interactivity (Whittaker et al., 1998). It is possible that given a large number of postings, short postings are read and replied to more frequently than long ones. Communication overload can also take place at the user-interface level of an email client (Whittaker & Sidner, 1996). Possible areas of breakdown in email clients include not showing threading, cluttered inboxes, inboxes containing hundreds of messages, and the diversity of information and cueing being shown within the UI. Coping strategies for dealing with communication overload are discussed by Whittaker and Sidner (1996). They describe how users develop workarounds for managing their email inbox, filing and finding information, and in general, handling email overload. Given that DLs use email clients for receiving, storing, and sending email, and that DLs are capable of delivering large volumes of emails, the functionality and usability of email clients is an important aspect of understanding how DLs are used, and potentially for understanding aspects of lurking.

Several researchers have talked about a volume of participants or critical mass required for the success of a community (Shenk, 1997; Jones, 1997; Grudin, 1995; Morris and Ogan, 1996) talk about a "critical mass" of members required to carry "free riders". In the case of DLs the critical mass would be a large enough group of posters to keep the public dialogue active, i.e., appropriate for the members needs. It is not known whether this is dependent or independent of the total number of members. There is evidence that key participants can make or break a DL. Berge (1998, personal communication) described a DL in which one person played a central role in a DL and when he left, the DL effectively died.

It is also likely that the volume of messages will have an impact on a DL (Jones, 1997; Morris & Ogan, 1996; Shenk, 1997). Whether a list is considered high in volume will vary from user to user. Among the factors affecting this determination are the number of messages received from the list, the number of messages received from other sources, quality of the messages, length of the messages, time available, the motivation for belonging to the list, and the email client and its usage.

DL membership can vary in number from two to hundreds of thousands (or more). If there is a relationship between size of group, posting levels, and number of lurkers, it has not been researched. Also, it is not known whether DL members know how many fellow members there are in a DL, or whether this is an important issue for them. One could certainly hypothesize that if members know there are many other members, then they may put less effort into posting, i.e., they would recognize that if all members posted

there would be anarchy. On the other hand, there may be more pressure in smaller lists for members to post.

In Roberts' study of the development of community newsgroups (Roberts, 1998), it was found that over two-thirds of the respondents had a sense of belonging and over half felt closeness within the group. For women, those with higher posting rates also had a greater sense of community. Roberts' results suggest that female lurkers should have a lower sense of community. Similarly, duration and frequency of posting have been found to be the best predictor for making friendships (Parks & Floyd, 1996). Their study suggests that lurkers should have low levels of friendship as lurkers' posting rates are low or non-existent. However, it is not clear whether lurkers who participate silently over longer periods of time, develop friendships through other means. Lurkers were not included in either of these studies.

Similar to the volume of messages, the topic of the DL and the number of topics/threads may have an effect on lurking. Whittaker et al. (1998) have used thread length as a measure of interactivity. The depth of the thread may also be important, e.g., long threads in DLs appearing to be linear are frequently multi-threaded conversations. Following these may take additional effort to read and follow. A very focused topic for a DL may also have an effect, as may the general type of the topic, e.g., would there be any lurking difference between health self-help groups and software self-help groups?

In their qualitative and quantitative study of the multi-user domain (MUD) LambdaMOO, Schiano and White (1998) suggest that design of virtual spaces should effectively support social interaction. If this does not happen, then people will find ways of circumventing the design. They indicate that designs need to support social interaction in the form of "private, personal spaces". This side-channel ability is built in to DLs in the form of private one-on-one email. While the lack of public participation is viewed as taking away from the good of the group (Kollock & Smith, 1996), little is known about the positive contribution mediated by either the built-in side channels or other user-developed side channels, e.g., face-to-face meetings, snail mail exchanges and telephone conversations. Parks and Floyd (1996) found evidence that these side channels were extensively used in the development of friendships over the Internet.

In summary, the five *related* questions developed in the first two sections of this chapter have guided the review of the literature in this section. In the process a rich picture has emerged of many important issues related to lurking. The *related* questions will be revisited in examining the results from the two studies (Chapter 5 & 6). The next section in this chapter examines the importance of studying online groups and DLs in particular.

2.4 The importance of studying online groups and DLs

This section reviews the literature on online groups, emphasizing why it is important to study both online groups and DLs. It begins with an introduction to DLs, and then splits the review of the pertinent literature into four parts, each showing the importance of studying online groups from a different perspective.

By way of introducing this section, the following is a short description of DLs. The DL is a simple and ubiquitous communication system that is quick to set-up and requires little maintenance. Email-based DLs have been commercially available since 1985 in pretty much their current form (and much earlier in non-commercial forms). With simple tools and information, i.e., an email account and the address of a DL server, an individual can subscribe to as many DLs as they wish. Depending on the size of the DL, a subscriber's email can communicate with a handful of members or tens of thousands of fellow subscribers. The well known receiving tool, the email client, the common server software, and the simplicity of the DL broadcast model are DLs' greatest assets.

Online groups and DLs specifically are becoming increasingly important areas of study. The following review of the literature examines this importance from four different perspectives on online groups:

- a ubiquitous and expanding technology
- room for change and improvement
- economic implications
- fertile gathering ground for multidisciplinary approaches

A ubiquitous and expanding technology: In their study of whether the Internet reduces social involvement and psychological well-being (Kraut, Mukhopadhyay, Szczypula, Kiesler, & Schelis, 1998a), Kraut et al. state that in 1998 40% of all US households owned a personal computer and approximately one in three of these homes had access to the Internet. They indicate that this rapid growth in the number of computers, coupled with the introduction of the Internet, is socially and economically transforming. In their study on the use of email in households, Kraut et al. (Kraut et al., 1998a) found that email drives people's use of the net. They state that knowing the importance of email has implications for engineering, policy development for the Internet, and for studying the social impact of new technology. As email is the basic message unit of DLs, their work suggests that the study of DLs is also important. Group communication tools similar to DLs have been examined and deemed important arenas of study. Whittaker et al. (1998) cite substantial growth in the number of newsgroups and view this area as a fertile ground for their work on *mass interaction*. They confirm there is little understanding of interaction in Usenet newsgroup. There is no reason to believe that DLs are any better understood.

LSOFT, one of several suppliers of commercial DL management software, advertises that there are over 55 million list members and over 150 thousand public and private lists using their software (L-Soft International, 1999c). These figures along with the extensive list of DL topics that can be found in the L-Soft DL catalogue, Catalist (L-Soft International, 1999a), indicates that DLs cover a broad range of topics and are widely used. As a means of group communication (although it often seems like individuals communicating publicly), DLs have gained wide acceptance, much of that resulting from using the most common of Internet tools, the email client. Email clients are also the basic receiving, viewing, and storage tools for DLs. While email is a novel technology for those new to the Internet, it has history spanning 30 plus years. DLs themselves have a history going back more than 25 years and in that time have gone virtually unchanged (Bennahum, 1996).

The same technology used in DLs is also used for the growing area of list publishing. List publishing differs from DLs in that an individual or small group is solely responsible for the distributed content, thus making the entire set of subscribers into enforced lurkers. A list of this type is not a public forum for dialogue, but a one-way broadcast from one to many, much like a paper-based magazine or newsletter. An examples of list publishing includes one of the longest running online magazine extant, Tidbits, a magazine for Macintosh computer users. In his Wired article Bennahum (1996), states "List publishing is not merely information delivered to your mailbox, it's the devolution of mass media into the hands of everyday people. And it's growing faster than the Web." Infrastructure start-up costs are low for this type of publishing, as are the costs associated with starting a DL.

Room for change and improvement: Kraut et al. (1998a) point out that change and improvement can come at many levels. They argue that the Internet's negative effects on consumers – they observed an increase in depression with Internet use – is not inevitable, but that changes in technology (design), deployment, and use by consumers will shape the effect. The emphasis on design, deployment and use echo a statement made by Winston Churchill 75 years ago when talking about great public buildings, such as the Houses of Parliament, and their effect on society:

> There is no doubt whatever about the influence of architecture and structure upon human character and action. We make our buildings and afterwards they make us. They regulate the course of our lives. (as quoted by Brand (1994))

It can be argued that much in the same way that public buildings shaped society in the past, communication technologies such as DLs, are and will shape the future. In their workshop report on the theory and practice of physical and network communities (Whittaker, Isaacs, & O'Day, 1997), Whittaker et al. call for increased insight into community technologies. They suggest that "community" systems offer different opportunities than those presented in current computer supported co-operative work (CSCW) applications and that it is "important that the field arrive at some insights into the theory and design principles associated with this novel class of system."

One of the most obvious areas where design, deployment and use can be studied and understood is the email client. For the DL member, this is the primary tool for accessing the DL. Because email clients preceded the development of DLs, email clients were designed for dealing with email from individual users, rather than groups. Nor is there evidence to suggest they were designed for dealing with high volumes of mail, or the ability to deal with the interactivity associated with DLs, e.g., viewing and relating multiple messages in a conversational thread.

There is little formal knowledge of how email-clients are used to deal with DL messages, although work has been done on how people manage their email in a specific email program in a business setting (Whittaker & Sidner, 1996). Wittaker and Sidner draw conclusions about how the tools are used and the strategies employed by users in dealing with "email overload". It is not clear from their work whether email overload is a result of users belonging to DLs, however, given the large number of emails/day (mean 49), this may be the case.

Economic implications: There are economic implications associated with online groups. Online groups can be related to specific products, e.g., CataList (L-Soft International, 1999a) shows 5 DLs related to the use of SPSS statistical software. These SPSS-related DLs are in effect, support groups for their respective communities. In their role as support groups, these DLs act as communication conduits amongst software users and between themselves and the developers of the software. Gripes, complaints and queries are often aired, as are official and non-official responses and announcements. They are rich environments for both the user and developer.

Sometimes the benefits are less direct for the sponsor of the DL. For example, Association of Cancer Online Resources, Inc. (ACOR) acts as host to 79 cancer related DLs that serve 39,387 subscribers. This provides a valuable resource to patients, their friends and families, and health-care professionals. The ACOR's About Us page (ACOR, 1999) describes their funding sources: "ACOR's activities are funded entirely by private donations and by service grants from major technology sponsors." It is unclear whether the private donations include private corporations who have a vested interest in the membership and content of the list. However, the technology sponsors are companies who provide hardware and software tools and support. Their involvement buys them goodwill and provides them with experience in serving health-support communities.

Other groups are imbedded in online enterprises, e.g., the Ask Dr. Weil Web site (Weil, 1999). This is a large site with a mix of information, community and advertising. For example, advertising banners have been built into the Web-based UI for each "community board" page. At the time of this writing, the top of each community board page contains a link to an index of advertisers, an ad for selling health related products such as vitamins, and an ad for non-health related items, e.g., Time magazine. Apart from the advertising at the top of each community board, the main Web site provides information and advertising, presumably balancing these in a way that is acceptable to the users of this site.

There are other economic avenues in which online groups can be used for financial gain. EBAY.com provides a venue where an association of sellers and buyers come together for mutual benefit. One of the ways in which the group dynamics play themselves out is in the feedback profile for each EBAY member. As members complete transactions, partners in the transaction provide feedback ratings for one another, either negative, neutral or positive. This history/profile becomes a source of reputation for both seller and buyers. For a buyer it may determine whether it is safe to bid on an item, i.e., is the seller trustworthy. For the seller, it has been used to prohibit bidders with poor records from bidding on items, e.g., notices such as the following are often found in an item's listing: "Cheques will not be accepted from bidders without at least a record of 10 positive transactions". In this community, one's persona is measured by the number of successful transactions and the lack of negative feedback.

Another economic opportunity lies in the email addresses associated with group members. All discussion groups where email addresses are used are subject to being raided for their membership lists. Anyone with access to the messages, whether as a subscriber or through accessing archives can access email addresses of members who post messages. Email addresses can be extracted from the communications and sold to anyone interested in directing advertising or other types of information at a specific group.

Not all economic implications are financial in nature. According to Bennahum's taxonomy of publication lists (Bennahum, 1996), lists can be separated by ownership into those that want to make money and those that don't.

> Typically, lists without financial motivation are animated by some other commitment, usually a desire to share in a collective enterprise – a political goal, the development of free software, the setting of public standards, the expansion of knowledge, or merely the pleasure of being heard, the joy of reaching other people and forming communities of shared interest.

He goes on to explain that remuneration does not have to be money, but can come as social status, fame, power, and secondary work such as consulting.

Fertile gathering ground for multidisciplinary approach: Not only is the Internet growing rapidly as a communication, entertainment, information and commerce tool, but the quantity of research has also been growing. Scientists with roots in communication, psychology, sociology have been attracted to this as an area of study. It has also attracted people from other areas such as design, political science, education, business and economics. As Morris and Ogan (1996) state, there is an opportunity to learn things about the Internet and its use that will illuminate our previously accepted understanding of traditional communication technologies.

While Morris and Ogan view this as a place of opportunity, others see it as an interesting problem space. Rafaeli, in discussing why the Internet should be studied by communication researchers (Newhagen & Rafaeli, 1996), suggests there is a research-engineering gulf: "The trick will be to think of concepts that bridge both the world of the engineer and that of the communication researcher. Two ways we might span those boundaries have to do with the technology's interface and its architecture." In a review of the literature, it quickly becomes obvious that there are two major camps: those of the researcher who's primary interest is in communication and the behaviour of groups (e.g., Roberts (1998) and Wellman(1997)), and the more engineer-like researchers interested in how technology is shaping this behaviour and what can be done to improve the technology (Whittaker and Sidner (1996) and Ackerman and Starr (1996)). It is not clear how these camps will come together.

Much like the field of human-computer interaction, which crosses many disciplines in order to understand how people use computers and how computers can be designed for people, the study of the Internet has and will continue to provide a basis for the marriage of different disciplines. The number of theories, models, and frameworks in this area is staggering, numbering in the dozens if not hundreds. Many of these refer back to earlier work, and are now being used across various disciplines. The Internet is becoming a test bed for previous understanding and also a means off creating new understanding. Part of its attractiveness is that it offers an environment in which monitoring of activity is relatively simple (if sometimes ethically unsound).

2.5 Chapter summary and direction forward

In summary, not much is known about lurkers even though the area of online groups is attracting significant attention from a wide variety of researchers. Lurkers are reportedly found at high levels in online groups although little is known about the variability between groups. There is general agreement that online groups and lurkers are worth studying, although it is clear that lurkers have not been studied. This is likely a result of the current methodologies focusing on the observable, i.e., message, and not lurkers and their non-posting activities. This review has outlined a number of areas where lurkers can be studied and where studying lurkers will have value.

In the next chapter, methods for studying lurkers in DLs are examined. Methods will be assessed for their ability to address the *primary* questions:

P1: Why do lurkers lurk?
P2: What do lurkers do?
P3: How many lurkers are there?

Chapter 3: Methodology review

Overview

- Six possible research methods for studying lurking are put forward.
- The methods are reviewed and assessed for their ability to address the *primary* questions.
- Two methods are chosen to study lurkers.

The previous chapter put forward three *primary* and five *related* questions about lurkers. In this chapter, six research methods used in the study of online groups are evaluated for their promise in addressing the *primary* questions. For each method there is a short introductory description. This is followed by lists of positive and negative aspects and a rating of the method's suitability in answering the three *primary* questions. This chapter is summarized with a discussion of the methods chosen for the two lurker studies (Chapters 4 and 6).

3.1 Possible research methods

The study of online groups has utilized many different approaches and levels of granularity. There are good reasons for this diversity. One of the most important is the complexity of the interactions being examined and the variety of questions being asked in the research. In his plenary talk at the 1992 Conference on Computer Supported Cooperative Work (CSCW), McGrath (1992) described research into the use of technology for collaborative work as a "complex puzzle". He went on to say:

> That both the theoretical ideas in this area and the available empirical evidence suggests that the effects of technology on collaborative work in groups involve very complex interactions of features of technology with attributes of members, features of group structure, type and characteristics of group tasks, and attributes of the context within which the group is working.

He might well have been talking about research into the Internet. Other researchers echo a similar point of view when it comes to understanding this relatively new area of research. Newhagen in his discussion with Rafaeli over the need for communication researchers to study the Internet (Newhagen & Rafaeli, 1996), discusses the failure of empiricists in providing an understanding of the strong effects of mass media. As he points out, this is largely due to the complexity of what is being studied. He suggests that the interactivity of the Net enables the successful study of these effects of mass media at the individual level and that researchers have to be comfortable with the study of interactivity across multiple levels of analyses. Both McGrath and Newhagen call for multiple approaches.

Kollock and Smith suggest that the Internet is a strategic place where fundamental social processes can be studied. "It provides a level of access to the details of social life and durability of the traces of social interaction that is unprecedented."(Kollock & Smith, 1999) However, researchers frequently look where the light is shining rather than where it is not . In the case of lurkers, the light has not been shining on them because their participation is generally not public and therefore not easily traced. Rafaeli and Sudweeks recognize how researchers' understanding has become lopsided because of it's focus on public contributions:

> When we come to the new reality of group CMC there is yet another split. That which is communicated, the messages, are the fruit of an unknown proportion of the participating audience. (Rafaeli & Sudweeks, 1997)

Researchers bring with them the heritage of their chosen fields and past research. A number of them have moved from a focus on CSCW to the Internet, e.g., Kraut and Wellman. Others have applied research traditions in other fields to the Internet, e.g., Preece (HCI), King (psychology), Rice (communication studies) and Erickson (design). This chapter looks at methods used in the study of online communities and groups with emphasis on implications for the study of lurkers in DLs.

Studies of online discussion groups generally use a number of methods (Anderson & Kanuka, 1997; Kraut, 1996; Kraut, Scherlis, Mukhopadhyay, Manning, & Kiesler, 1996; Schiano & White, 1998; Whittaker & Sidner, 1996). Multiple methods are called for as no single method is capable of providing all the information or answering all the questions. In addition, multiple methods often provide redundancy and validity checks (Thomsen, Straubhaar, & Bolyard, 1998).

Some methods, such as demographic surveys are useful in providing population information in the form of patterns of activity, but fail to provide answers to why the patterns occur. Other methods such as observation provide a contextually rich view of people and their work, but are time consuming. Many of the techniques listed below can be used face-to-face, over the telephone, or online. Some may also be used either in a stand-alone fashion or nested within one another. For example, an online questionnaire can be used on its own or as part of an interview. The following is a list of six methods that have shown value in studying online groups:

- logging
- questionnaire
- interview
- observation
- ethnography
- content & discourse analyses

Each method is briefly described and reviewed below. Positive and negative aspects as related to the study of lurkers are listed. The method is then rated for each of the *primary* questions. The rating is approximate and reflects the time, expense, expertise required and effort. The rating for each method also incorporates its perceived its ability to address the *primary* questions. The *primary* questions are:

P1: Why do lurkers lurk?
1. P2: What do lurkers do?
P3: How many lurkers are there?

The rating scale has three levels: not recommended (-), some potential (?), and recommended (+). The chapter concludes with a discussion of the two methods chosen to study lurkers.

3.2 Review of six methods

Logging: Software logging is the process of gathering records of behaviours and activities. In the case of DLs, a useful logging event is the distribution of outgoing email from the list. The artifact is the email, which includes headers and content. Message logging can be accomplished by gathering messages over a period of time or by using archives or other repositories. Logging studies range from the focused study of a DL-based community of journalists (Millen, 1997; Millen & Dray, 1999) to the study of mass interaction in a broad cross section of newsgroups (Whittaker et al., 1998). A typical logging operation consists of collecting messages and then performing an analysis, frequently by counting, and/or sorting using various criteria. Logging has been used to examine many issues including Internet usage in households (Kraut et al., 1998a) and conversational strategies in newsgroups (Whittaker et al., 1998).

It should be noted that newsgroups differ from DLs in a very important way. There is no way through the use of software to determine the number of members receiving messages in a newsgroup. In DLs, the number of members can be determined by querying the DL server. Thus, Wittaker et al.'s (1998) data on

newsgroups could not be used to determine lurking levels. Other means such as questionnaires would need to be employed to determine lurking levels in newsgroups.

Whittaker et al. (1998) have shown that log-based surveys are excellent at not only pointing out interesting issues directly, but highlighting areas worthy of further investigation through other techniques and studies. As a preliminary tool, they have been used as mechanisms to select individuals to be surveyed through questionnaires (Roberts, 1998). Message logs can be used as the raw data for demographic surveys, and also as the basis for discourse or content analyses (Preece & Ghozati, 1998). Their multiple utility (Smith, 2000; Smith, Farnham, & Drucker, 2000) makes them particularly valuable when studying systems such as DLs.

The negative and positive aspects of this method are outlined next. This is followed by a discussion on the value of using this method for studying lurkers and concludes with a cost-benefit rating of the method for examining each of the *primary* questions. This structure is used in describing each of the research methods.

Positive aspects

- easily automated way of collecting large amounts of data
- can collect from multiple sources simultaneously
- unobtrusive asynchronous data collection
- information is often public, e.g., DL archive and/or messages by subscription
- useful in finding out about usage/activity patterns
- provides a large number of opportunities for quantitative analyses
- often points out further areas to be studied
- data can be analyzed from many different perspectives

Negative aspects

- raw data currently requires customized tools to parse into analyzable format
- choice of group and sampling method is critical
- danger of being swamped with data and choices for analyses
- can create very large volumes of data, e.g., 2.15 million messages (Whittaker et al., 1998)
- doesn't answer the question "why" very well, i.e., provides numbers, not explanations

Value for studying lurking

This may be the only cost effective way to determine lurking levels across a large number of groups. By examining those who post, and then comparing this to the number of members there are in a given group, it is possible to find out how many members are lurkers. By logging messages in a number of different types of DLs, lurking levels can be compared between different types of communities. This technique has the added advantage of being able to capture other information that may be relevant to lurking levels, e.g., volume of messages, size of messages, message content, and threading of messages. The data captured through logging can also be used for other types of investigations, such as the study of flaming through discourse analysis.

As noted above, logging is a very useful way of gaining quantitative data about a population. It is of no direct value in determining how lurkers lurk as this is unobservable when logging emails from a DL. To find out how individuals interact with DL messages, logging of email usage at the client level would be required. This would provide a better picture of tool usage, but would fail to provide reasons for the lurking. It would also be a complex proposition as each client would need to be instrumented. In either case, it would be difficult to attribute actual activity from logs. For example, just because someone opens an email, does not mean they have read it. Logging receives a "some potential" rating for P1 because logging data can be used to determine if correlations exist between lurking levels and other factors, e.g., traffic volume. These correlations will not indicate the cause of the relationship, only whether it exists and to what degree.

Logging: Cost-benefit rating	Research Questions		
	P1. Why do lurkers lurk?	P2. What do lurkers do?	P3. How many lurkers are there?
	?	-	+

Note: - (not recommended) ? (some potential) + (recommended).

Questionnaires: Questionnaires are a time tested means of gathering information and have been used in many studies. They have been used for many different purposes ranging from determining the nature of online friendship (Parks & Floyd, 1996) to measuring the extent to which people experience community online (Roberts, 1998). They have many forms, ranging from short email delivered questionnaires (Mason, 1999) to comprehensive multi-level surveys (Garton, Haythornthwaite, & Wellman, 1997). Questionnaires can be used to gather qualitative or quantitative data and can be administered in many different ways, e.g. from face-to-face interviews to Web-based questionnaires (Lazar & Preece, 1999). They can be used as preliminary tools in selecting interview subjects and/or be a part of an interview process (Preece, 2000; Preece et al., 1994).

Positive aspects

- well developed methodology that is used extensively, e.g., questionnaire for user interaction satisfaction, QUIS (Harper & Norman, 1993)
- relatively easy to deploy on the Internet, i.e., Web interface and email
- data can be compiled automatically
- no geographic barriers if carried out over the Internet
- can be qualitative or quantitative or a mix
- closed and open ended questions are possible
- can be used to examine differences over time

Negative aspects

- often used in situations where participants self select; this can lead to unintended biases in results
- very low response rates from lurkers (Mason, 1999)
- response rates from non-lurkers are typically less than 20% (Mason, 1999; Schiano & White, 1998; Smith, 1997)
- difficult to do immediate follow-up, unless administered in an interview situation
- if the question "why" is not asked, results can be difficult to interpret
- participants may perceive questionnaire as tedious, especially if not well designed

Value for studying lurking

A basic problem with questionnaires is the difficulty of obtaining information that is not biased by the population. Given the nature of lurking, lurkers may be less likely to respond to online surveys. Evidence indicates that lurkers respond to questionnaires (Mason, 1999) at a much lower rate than public posters (in Mason's study, ~3% of responses were from lurkers and 90% of the membership lurked), and it may be difficult to understand why those that do respond, do so. This may result in a biased view of lurking. As a means of structuring interviews, questionnaires have value as it allows some semblance of equivalent information taking between participants. However, questionnaires, if followed religiously, can act to occlude important information and follow-up.

Questionnaires are a basic tool for providing both quantitative or qualitative results. They can be used in many situations and there are plenty of examples of their use. The difficulty in using them to study lurkers, specifically in online situations, makes their use in this manner somewhat questionable. However, if they

can be used where the population has been selected in a balanced way, then their value would improve. The ratings below reflect their reliance on selection technique. Questionnaires, when used on their own, often fail to provide answers to "why" type questions. At the current stage of understanding lurking, it is expected that many "why" type questions will arise.

Questionnaire: Cost-benefit rating	**Research Questions**		
	P1. Why do lurkers lurk?	P2. What do lurkers do?	P3. How many lurkers are there?
	?	?	-

Note: - (not recommended) ? (some potential) + (recommended).

Interview: Interviews provide a mechanism of talking with participants and gaining insight into their practices and issues. Explanation can be gathered from the participant without the process becoming burdensome to them (as can extensive questionnaires which take considerable time to complete). Interviews range from open-ended, where the process of discovery is a primary goal (Barry, 1995; Schiano & White, 1998), to structured interviews where a questionnaire is delivered by the interviewer. Interviews can be carried out in a number of ways, e.g., face to face, over the phone, or via the Internet through synchronous tools like chat or asynchronously via email. There is evidence that for questionnaires where anonymity is preserved, participants may be more forthcoming (Hewson, Laurent, & Vogel, 1996).

Positive aspects

- allows follow-up where questions of type "why" can be asked
- can provide very rich data, i.e., good interviews are frequently more than just a verbal response to a questionnaire

Negative aspects

- requires experienced interviewer and interpreter
- requires effort to pre-select candidates ahead of time
- can be very expensive if participants not readily available and/or travel required
- in high volume, can take much time and effort
- analysis can be slow and labour intensive
- participants may rationalize their actions

Value for studying lurking

Interviews, especially with small populations, are an effective means of jump starting an investigation. As mentioned above, they can be structured using a questionnaire. They are excellent tools for obtaining peoples' stories, and for providing a flexible information gathering process where the question "why" can be asked. The open ended interview is a powerful tool for developing both a broad and deep sense of the problem space (in the case of lurkers: why they lurk and what they do). Ethnographers often use interviews as one of a set of information gathering tools (as described further on in the review of ethnography). Without proper sampling of a population, biased results will result. If proper sampling is not possible, then understanding the biases is necessary in order to interpret the results.

Interviews allow the participant to be asked directly about what they do and why they do it. Interviews allow interaction between interviewer and participant in ways that are difficult to achieve in any other way. In the case of lurking, where the knowledge of lurking is essentially zero, interviews will be useful in understanding the variation between participants and the reasons for this variation. Interviews are excellent

ways of gaining insight into novel domains. Questionnaires can be used within an interview to provide quantitative information. For example, questions of the type: How many online groups do you lurk in?, can be asked. For this reason, the interview method receives a "some potential" rating for addressing P3: How many lurkers are there.

Interview: Cost-benefit rating	**Research Questions**		
	P1. Why do lurkers lurk?	P2. What do lurkers do?	P3. How many lurkers are there?
	+	+	?

Note: - (not recommended) ? (some potential) + (recommended).

Observation: Observation is perhaps the most sensitive method to the context of the user and as such, is excellent for getting a rich picture of the impact of their environment, the activities taken, and the strategies used. However, it is perhaps the most demanding from the researchers perspective as it requires an observant and attentive researcher. It has been used for many different types of work, from examining how people use photocopiers (Suchman, 1987), to large scale commercial software projects (Holtzblat & Beyer, 1995). Observation comes in many flavours, from very quick turnaround studies where specific issues are examined, to intensive examinations of information and artifacts used (Macaulay, 1999). Observation can be carried out under "controlled" conditions in a lab or in a participant's own work or home environment.

Positive aspects

- brings forward the context of the activity
- often complimented with interviews or questionnaires
- very rich data

Negative aspects

- difficult when activity is dangerous or of a low frequency (e.g., observing nuclear melt downs)
- can be a costly, intensive process
- often requires substantial time on the part of researchers and participants
- can intrude on participant

Value for studying lurking

Observation is perhaps the best tool for understanding the context of an activity. It is also one of the most expensive and time consuming processes. It can also be one of the most intrusive processes, as it frequently involves the observer (sometimes a participant observer) looking over the shoulder of another person. In the case of online observations, messages frequently become the observed activity. Online ethnographers frequently read messages of online groups in order to understand the community (e.g., Mason, 1999). It is a difficult process when events occur at low rates or where there are long intervals between events. Once again, the nature of the lurker may preclude this technique from being used, both because of the low frequency of the events, and also because of the apparent nature of lurking, i.e., privacy may be an issue.

Unlike message logging, where events are recorded with little context, observation is a type of event recording in which context can be understood. Direct observation of lurkers would be useful in understanding how individuals use their tools while lurking. For example, an observer could watch how messages are manipulated in the process of lurking. It would also provide the opportunity for understanding why they do what they do, as questions could be asked to draw out this information.

The great difficulty with direct observation is that the activity of lurking is not well understood. It may be an infrequent activity, and observing may be very time consuming relative to the value received. Observation's primary value lies in understanding activity in context, and not in providing demographic type information as required for P3. Observation will be invaluable in understanding how tools such as email clients are used while lurking.

Observation: Cost-benefit rating	**Research Questions**		
	P1. Why do lurkers lurk?	P2. What do lurkers do?	P3. How many lurkers are there?
	-	?	-

Note: - (not recommended) ? (some potential) + (recommended).

Ethnography: Ethnography is not a method per se, but an approach. It is a decoding operation in which a shared knowledge of cognition allows the researcher to decode the observed behaviour (Thomsen et al., 1998). It is carried out using observation, collection of written materials and artifacts, interviews, insider information, and participation in the community (Denzin & Lincoln, 1994; Walsham, 1993). The word immersion is often used in describing ethnographers' work (Fetterman, 1998) and the process of immersion is frequently described as participant observation, i.e., the observer becomes a participant.

With the realization that the Internet is a fertile ground for researchers, lurkers may be researchers as Thomsen (1998, p. 11) describes: "It is much easier to lurk on the Internet in most cases than to unobtrusively hang out in an Amazon village". Mason (1999) also talks of virtual communities as being an "ethnographer's paradise: a way to observe without being observed", but cautions, that by only observing in the public spaces everyday behaviour is missed behind the closed doors. He argues that ethnographers will have a richer experience by becoming public participants in the community. Similarly, in the study of LPMUD, Karetnick's took the role of participant ethnographer (Karetnick, 1998).

Mason (1999, P.62) states that in order "to study the virtual community, populated by virtual people then we must become virtual ethnographers." He goes on to describe three basic strategies for studying a DL based community. The first is to join a DL and read the messages as they occur or through archives, and participate as an ethnographer, "by asking questions, contributing to debates, sparking conversation and so on". He indicates that not all communication within a DL-based community occurs in the public space. The second is to perform electronic surveys. This provides insight into the people within the DL, but as Mason warns, the lurkers, who constitute 90% of the DL population he studied, responded at very low levels. And third, perform email interviews with specific questions, where the interview process and interviewer are obviously identified as such.

Variations on ethnography have been applied for different purposes in a range of contexts. For example the study of social networks through questionnaires (Garton et al., 1997); the in situ study of multi-media publishing (Bellotti & Rogers, 1997); the study of control room usage (Hughes, King, Roden, & Andersen, 1994); the study of information systems use by researchers (Barry, 1995; Barry, 1997; Squires, 1997); the study of MOOs (Karetnick, 1998); and the study of a specific DL (Mason, 1999).

Positive aspects

- results depend on skill of participant observer
- ethnographer can immerse him/herself in the group
- good at developing longitudinal knowledge
- rich/thick data
- provides tremendous insight into the communication patterns of the group and the individuals
- shown to be useful for studying a DL

Negative aspects

- depends on skill of participant observer
- requires commitment, often long term
- frequently takes a long time to perform
- validity and authority can come into question
- results can be difficult to translate into design (Dillon (1998) discusses this in detail)
- results may not be able to generalize beyond particular group
- often done too late in the (software) design process; ethnographers are rarely designers
- reporting can be difficult

Value for studying lurking

This is an intensive process that has high value in studying social networks and has shown to be of great value in studying a particular community. It has less value in studying multiple communities due to the requirement that the participant observer must belong to many communities. This is possible, but difficult when the number of communities range in the tens or hundreds.

Without limitations, such as the budgets and the number of researchers, ethnography's broad use of the methods mentioned so far makes it an ideal candidate for studying both the reasons for lurking and also the activities of lurking. Over a period of time, it could also provide the quantitative data associated with trying to understand how many lurkers are out there. Whether it is the best candidate for researching lurkers is largely up to external limitations, e.g., time available. This is a case where a methodology, which is capable of producing excellent results, is down-graded to "some potential". This primarily due to the amount of time it takes to do an ethnography and the limited time available.

Ethnography: Cost-benefit rating	**Research Questions**		
	P1. Why do lurkers lurk?	P2. What do lurkers do?	P3. How many lurkers are there?
	?	?	-

Note: - (not recommended) ? (some potential) + (recommended).

Content and Discourse Analyses: Content analysis is the analysis of the content of the messages, while discourse analysis follows the exchange and relationships between messages. While of less interest in this thesis, content and discourse analysis has been used to understand the nature of the message within self-help groups (Preece, 1998; Preece & Ghozati, 1998; Winzelberg, 1997) and in the general study of interactivity (Rafaeli & Sudweeks, 1997). A possible use in studying lurkers in DLs would be in understanding whether certain types of content or interaction are related to lurking, e.g., do DLs with high flame rates also have higher lurking rates?

Positive aspects

- provides statistical reliability and validity
- developing categories provides a substantial amount of insight
- qualitative and quantitative results are possible

Negative aspects

- categorization tends to lessen the value of researchers' insights
- analysis is often time consuming and expensive
- inter-researcher reliability can be difficult to obtain
- developing categorizations can be a time consuming and difficult process

Value for studying lurking

This technique is excellent, if expensive and time consuming, for understanding the nature of interaction. Unfortunately, lurkers do not interact in the public spaces, and it would be all but impossible to track side posting in a DL. In other group tools, such as closed systems used in businesses, side posting could be tracked, however, the ethics of doing so would have to be considered. One area where this tool could be used is in the analysis of people who delurk, i.e., reveal themselves for the first time. Understanding how they present themselves, their reasons for doing so, and how they are responded to by the group may be very revealing about both the lurker, and the nature of the group.

Lurkers, being partially defined from their absence in public spaces, are a difficult group to study directly through content and discourse analyses. However, combined with other methods, such as logging, where levels of lurking can be established, content and discourse analyses may prove to be valuable in understanding why lurking levels vary, and why lurkers lurk. Content and discourse analyses could provide value in understanding the use of side channels, such as a direct exchange of emails between members of a DL. In a sense, Katz's understanding of lurkers has come from this type of exchange, although not through a rigorous analysis.

Content and Discourse Analyses: Cost-benefit rating	**Research Questions**		
	P1. Why do lurkers lurk?	P2. What do lurkers do?	P3. How many lurkers are there?
	-	-	-

Note: - (not recommended) ? (some potential) + (recommended).

3.3 Methods chosen to study lurkers

The study of lurkers and lurking is largely untapped and the explanations and observations so far put forward are not well supported by research. Specific views have been put forward, e.g., lurker as free-rider (Kollock and Smith, 1996), but it is uncertain whether these are the "right" views or promote the right questions. What is needed is an inductive process, one that informs the research community about lurking and also provides focus for further work. It is for lurkers to inform researchers about their lurking and not for researchers to impose a perspective on lurking.

The choice of how to study lurking is dependent on the questions being asked, the techniques available, and any constraints associated with the research. The *primary* questions being asked, as developed in Chapter 2 are as follows:

P1: Why do lurkers lurk?
P2: What do lurkers do?
P3: How many lurkers are there?

All of the methods described in this chapter could be used to uncover answers to the primary research questions, however, as Table 3.1 shows, some are better suited than others. At first glance, the interview and logging methods appear to be the best choices. Understanding how the constraints of the research affect each of these methods partially determines which of these methods will be most fruitful in providing answers to the research questions.

Research method	**Research questions**		
	P1. Why do lurkers lurk?	P2. What do lurkers do?	P3. How many lurkers are there?
logging	?	-	+
questionnaire	?	?	?
interview	+	+	?
observation	-	?	-
ethnography	?	?	-
content & discourse analysis	-	-	-

Note: - (not recommended) ? (some potential) + (recommended).

Table 3.1: Methods rated for effectiveness in answering the three primary research questions.

Constraints

There are a number of constraints that need to be considered. The chosen methods must complement one another, providing both an initial broad understanding and uncovering further avenues of research. The techniques should be as independent of one another as possible, reducing the possibility of an interdependence disaster in case of an early failure. The initial technique should be productive as soon as possible and any other method should be able to be carried out in parallel. The results must cover not only the qualitative results, but also provide a quantitative picture of lurking. Sudweek and Simoff (1999) describe this approach in their work on complementary explorative data analyses. Sometimes called *triangulation* or *mixed method*, the combined quantitative and qualitative approach "results in an integrated view that narrowly focuses on a particular social phenomena." (p. 37).

At an operational level there is another set of constraints, one which reflects the nature and circumstance under which the research will be carried out. These include completing the research within a reasonable time frame, having one individual perform all research functions, and taking advantage of resources such as readily available participants and Internet facilities. The methods must be relatively quick and

straightforward to implement, and low in cost. This means they must not require overly specialized programming skills or skills that are not readily available within the local research community. In addition, one of the aims of this work is to show the value of the chosen methods and component techniques.

There are ethical considerations that must be met. The methods should retain the privacy of the participating individuals. Reported information should not identify individuals and their anonymity should be preserved. This is also true of any messages collected from DLs. While DL messages are generally archived in some public form, it is not clear whether DL members realize this. Given the sophistication of current search technologies, quoting information from messages could lead to the identification of the messages' authors. If quoting is of value, it will be necessary to gain the consent of the author of the message. The choice of DLs should be in compliance with the rules of that DL. For example, if a health-support DL requires members to have a specific affliction, then information should not be collected from that DL. If personal contact is made with participants, e.g., for an interview, then full disclosure as to the purpose of the research and the affiliation of the researcher should be made.

If time were not an issue, ethnography would be a preferred method for studying lurkers. Given the constraint that the research must be completed in a reasonable time, the interview research method was chosen. Interviews can provide a broad picture of lurking without the long term commitment associated with ethnography. It was reasoned that because so little is known about lurking, that open-ended interviews would provide the most efficient way of uncovering the greatest amount of information in the shortest period of time. At this point in the research, a broad picture is perhaps more valuable than a deep picture. As interviews are often a component of ethnography (sometimes called the ethnographic interview), it was reasoned that these interviews could become the basis for ethnographic studies at a later date. As for counting the lurkers, Table 3.1 shows logging to be the favoured method. In addition, it is a well understood technique having been successfully used in previous studies of online documentation use by the author (Hendry et al., 1990; Nonnecke, 1992) and by others in studying online social structure (Smith, 1999). Both interview and logging methods have the ability to compliment each other. For example, interviews may indicate that lurking occurs more frequently in DLs with high traffic rates. This could then be verified with the logging data.

3.4 Chapters to follow

Based on the constraints and the suitability of each method to provide answers to the research questions, a parallel approach will be taken. Two studies will be performed each with a different method (see Table 3.2). The first will use open-ended interviews loosely structured around a questionnaire to uncover as many issues as possible. It is expected that this study will provide insight into why lurkers lurk and what they do. The second study will examine lurking from a quantitative perspective using message logging across a number DLs over a several month period. The results from the second study will indicate whether there are relationships between lurking and community size, type, and volume of messages. The combination of these two independent studies should produce a valuable profile of lurking at the individual and group levels. The methods used in these studies will be described in detail in Chapters 4 and 6. Chapter 5 contains an in-depth discussion of the interview results presented in Chapter 4.

Research method	Research questions		
	P1. Why do lurkers lurk?	P2. What do lurkers do?	P3. How many lurkers are there?
logging	?	-	+
interview	+	+	?
Chapter & Study	**Chapter 4:** Lurkers speak – Interview & results **Chapter 5:** Discussion of interview results	**Chapter 4:** Lurkers speak – Interview & results **Chapter 5:** Discussion of interview results	**Chapter 6**: Demography results: Counting the lurkers

Note: - (not recommended) ? (some potential) + (recommended).

Table 3.2: The two research methods to be used in studying lurkers.

Chapter 4: Lurkers speak - Interviews & results

Overview

- The interview method is described.
- An overview of the results is presented.
- Results are presented related to *P1: Why do lurkers lurk?*
- Results are presented related to *P2: What do lurkers do?*
- Results are summarized.

As described at the end of the previous chapter, two studies utilizing different investigative methods will be used to study lurkers and lurking. This chapter contains a description of the interview method used in the first study and the interview results. The findings are split into three sections: an overview of the results and two sections, each addressing one of the *primary* questions:

P1: Why do lurkers lurk?
P2: What do lurkers do?

The chapter concludes with a short summary. A discussion of the results in this chapter can be found in Chapter 5. The second investigation, a message logging study examining the *primary* question, *P3: How many lurkers are there?*, can be found in Chapter 6.

4.1 Interview method

The goal of this work is to develop a preliminary understanding of lurkers and lurking. The study should reveal as much as possible about lurking in as rich a way as possible. To this end, interviews with online group members were chosen as the primary method in finding out why lurkers lurk and what they do while lurking.

Participants were selected at random from two physical communities in which members were known to be Internet users. Given the relatively high incidence of lurkers reported by Mason (1999), it was assumed that the majority of participants would more than likely be lurkers. Ten interviewees were drawn from two locales, 5 men and 5 women, ranging in age from early 20s to early 50s. The intention with the small sample size was to balance for age and gender, rather than examine age or gender issues. All participants were members of at least one online group, and were not pre-selected for lurking or for their level of experience with online communities. All persons asked participated in the study; 3 were well known to the researchers, and 7 were not.

Face-to-face or phone-based, open-ended interviews lasted between 45 minutes and 2 hours, and focused on the interviewee's participation in online groups. Prompting was minimal, and the interviewer did not validate whether a group or topic was worth discussing. The following outline was used to guide the interviews:

1. Introduction
2. Name and describe online groups to which they belong(ed)
3. For each group, determine the following:
   - reasons for joining
   - activities and action they took (in posting or non-posting)
   - reasons for not-posting (if this occurred)
   - reasons for posting (if this occurred)
   - reasons for side posting in non-public spaces (if this occurred)
   - reasons for leaving the group (if this occurred)
4. Ask for comments on Jargon Dictionary definition
5. Ask for any additional comments
6. Thank them for participating

4.2 Overview of results

The participants described 41 groups of which there were 25 DLs, 7 BBSs, 5 newsgroups, 3 chat rooms, and 1 MOO (see Table 4.1). All participants belonged to or had belonged to groups in which they never posted, or posted rarely, e.g., once or twice, or so infrequently that they considered themselves to be lurkers. Two interviewees belonged to only one group while the largest number of groups mentioned by an individual was 8. This member mentioned 6 DLs. All participants had posted in at least one of their online groups, even if they did consider themselves to be lurkers in those groups.

	Group Type					
Participant	DL	BBS	newsgroup	chat	MOO	Lurked/Total
P1	LLL		LLL			6/6
P2	LLLL					4/4
P3	L					1/1
P4	LLLLLL		L		P	7/8
P5	LLLL-PP					4/6
P6	L	L-P				2/3
P7	L-PP	L				2/4
P8		LL		PP		2/4
P9	L	P	P	P		1/4
P10		L				1/1
Lurked/Total	21/25	5/7	4/5	0/3	0/1	30/41

Note: L=lurked in a group P=posted in a group

Table 4.1: Group types and whether lurking occurred.

The definition of lurker used to construct Table 4.1 is a simplification of the definition found in the Jargon Dictionary. In this case, lurker is defined as anyone who for prolonged periods receives communications without posting.

While the sample is small, it is interesting to note that the asynchronous groups (DL, BBS, and newsgroup) have lurking rates around 75%, the synchronous environments (chat and MOO) have no lurkers. As one

participant (P4) pointed out, it is much more difficult to lurk in synchronous environments as you are almost always immediately visible to other members. Being visible acts as an invitation for others to approach. This participant went on to indicate, that the reason for entering these places is to converse with others. This participant lurked in all DLs and a newsgroup in which she belonged.

To provide a flavor of the participants' lurking, two participants' experiences are presented. The first describes the impact of receiving messages in digest form and the second describes the general lurking process of a participant.

Cathleen's lurking using digests. Cathleen is a well known member of high standing in a professional organization. She is also a very private person. Having a health problem, she sought out and subscribed to an online DL specializing in her health problem. She read and saved all messages in digest form for several months, but found it difficult to follow threaded conversations. Some digests were printed to aid reading and to save important messages. After the initial period Cathleen started skimming the digests, reading individual posts based on the subject heading. As her health improved, the DL became less important to her. Throughout her membership, she found the moderation to be aggressive and disruptive. She was eventually removed from the DL by a moderator who falsely accused her of cross-posting.

Fred's general lurking strategy. Fred is a knowledgeable, long-term user of DLs and is a technically sophisticated group member. He has belonged to a variety of DLs for reasons ranging from professional to personal interest. As a general rule, he does not post to DLs, preferring to post directly to individuals based on their public posts. He belongs to personal interest DLs to learn about the communities and for entertainment.

Fred follows threads but does not read every message in the thread. If he is very busy, he will delete messages without reading them, confident that the same issue will arise at a later date. When investigating a particular message, he uses the subject header and reads the first paragraph before continuing on. He reads to discover others' problems (e.g., technical problems with software), and says it is difficult to find this type of information in any other way, i.e., it is hard to ask about a specific problem when you don't know the problem exists.

He is also interested in learning about the community, stating that learning about the members helps him to learn about the community. He systematically described his method of coming to know members:

- information is gleaned from email address, name, signature, and URL
- understanding the members comes from what each says and how it is said
- inferences can be drawn from the choice of a false name
- knowledge about posters' habits comes from their frequency of posting and the time of day they post.

This list is very similar to the one Donath has put forth in her work on how individuals show their identity in online groups (Donath, 1996). On joining a new DL, Fred reads every message to get a broad sense of the DL. He looks for cross-posts as they tell him how members view the DL in the context of related DLs and newsgroups. DL rules describing topics, moderation policy, and membership requirements, etc. tell him much about the community. Likewise, what members say and how they say it is also informative.

The Jargon Dictionary indicates that lurkers read posts on a regular basis and post occasionally or not at all. In contrast, the next section describes why the interviewees lurked.

4.3 P1: Why do lurkers lurk?

When asked why they lurked, interviewees provided 117 reasons. These have been grouped into 33 categories and are presented in seven tables each containing a related set of reasons for lurking:

- personal (Table 4.2)
- relationship to group (Table 4.3)
- intention from outset (Table 4.4)
- wants/needs met by lurking (Table 4.5)
- group characteristics (Table 4.6)
- stage of membership (Table 4.7)
- external constraints (Table 4.8)

After each table is introduced, the table is presented and a description of the contents is provided.

Table 4.2 lists reasons for lurking which were personal in nature. The reasons generally reflect a desire to reduce public exposure.

	Participant										
Personal	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total
2. remain anonymous, preserve privacy and safety	*	*	*			*	*	*	*	*	8
3. shy	*	*					*	*		*	5
4. can't offend					*	*					2
5. difficulties with language								*			1
Totals	2	2	1		1	2	2	3	1	2	16

Table 4.2: Personal reasons for lurking as cited by participants.

1. Remain anonymous, preserve privacy and safety: Eight of the ten participants cited these as reasons to lurk. Half of the participants indicated they wanted to remain anonymous. Reasons for wanting anonymity varied between participants. For example, two participants wished to remain anonymous because they did not want their professional position to be known to their groups. Six of the ten participants cited privacy and safety issues. Many of these participants understood that messages were persistent and that they would be available to others for an indefinite period of time. In one instance, the participant recognized that posting might bring unwanted attention from the opposite sex. Two of the participants were aware that messages posted to public groups were a source of addresses for junk email, i.e., spam.

2.Shy: Akin to a fear of public speaking, half the participants cited shyness and discomfort about public posting as a reason for lurking.

3. Can't offend: Two participants cited lurking as a way of belonging to a group and at the same time limit their ability to offend others in the group. Both of these participants belonged to groups in which tempers were frequently expressed and opinions were freely given. Flaming by members in these groups was common and participants in this study did not want to directly feel the heat.

4. Difficulties with language: One person was a non native-English speaker and was concerned with how his written English would appear to others.

Table 4.3 describes how individuals see themselves in relation to the group, either socially or from an informational perspective. Their relationship to the group either inhibits their public participation or reduces the need to participate.

	Participant										
Relationship to group	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total
1. not part of the group		*	*			*			*		4
2. nothing to offer		*		*	*	*					4
3. lack expertise	*			*							2
4. know it will be answered		*			*						2
Totals	1	3	1	2	2	2			1		12

Table 4.3: Relationship to group cited as reasons for lurking by participants.

1. Not part of the group: Four of the participants saw themselves as being distinct from the group and until they somehow become full-fledged members of the group, the distinction is a barrier to public participation.

2. Nothing to offer: Four indicated that they had nothing to offer. This appeared to be a case in which there was some kind of mismatch between the participant and the list. This mismatch could be the result of the topic not being of interest or as in one case where several DLs were a mandated part of their work.

3. Lack expertise: Several of the participants indicated that a lack of expertise in the topic area was a reason for their lurking. One indicated that it was much easier to ask a question than reply to one. The "lack expertise" reason suggests that the group is some sort of informational font, and that the participants were unable to add value.

4. Know it will be answered: Two of the participants indicated they knew others in the group would respond. This absolved them of the responsibility of replying and also reduced duplicate responses.

Table 4.4 lists reasons why participants lurked from the outset. Seven of the ten participants intended to lurk from the outset and gave a variety of reasons for doing so.

	Participant										
Intention from outset	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total
1. never intended to post	*	*		*			*				4
2. no specific need to post	*					*		*			3
3. not motivated to post			*			*					2
4. reduce involvement/commitment	*	*									2
Totals	3	2	1	1		2	1	1			11

Table 4.4: Intentions from outset cited as reasons for lurking by participants.

1. Never intended to post: For four of the participants, there was never any intention of posting when joining the group. This was decided upon even before joining the group.

2. No specific need to post: Not having a specific need to post was also cited as a reason to lurk. These participants saw posting as showing a need for information and if no information was required, they did not post.

3. Not motivated to post: Two others indicated that there was no motivation to post. In one case this was a result of there being so much content that the participant used his/her time for reading and not posting. In the other situation, the participant was not about to be drawn into the group conversation.

4. Reduce involvement/commitment: Two of the participants indicated that when they lurked, they were making less of a commitment to the group. This was important to them as this meant they had less of a problem leaving the group, i.e., they made no commitment so leaving was not breaking a commitment.

Table 4.5 lists the reasons given for lurking related to meeting their wants or needs. For all of the participants, lurking provided a means of satisfying their wants or needs. This is not to imply that lurking satisfied all their needs, only that it was capable of supplying a variety of wants/needs.

	Participant										
Lurking provided	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total
1. conversation/stories	*	*	*	*	*	*	*	*			8
2. entertainment	*	*		*	*			*	*	*	7
3. information, not interaction	*		*		*	*	*	*			6
4. access to expertise/experience	*	*	*	*			*				5
5. community without posting		*		*		*	*	*			5
6. connections with individuals	*	*		*			*				4
7. stuff in mailbox				*	*				*		3
Totals	5	5	3	6	4	3	5	4	2	1	38

Table 4.5: Wants/needs met by lurking cited as reasons for lurking by participants.

1. Conversation/story: Eight of the ten participants lurked in order to follow conversations or hear others' stories. In this case, stories were synonymous with the personal experience of the posters and considered valuable.

2. Entertainment: Seven of the ten participants said they lurked as a form of entertainment. The entertainment came in many forms including flames (without getting burnt). Additional sources of entertainment came in the form of humorous messages and jokes.

3. Information, not interaction: Lurking was seen by 6 of the participants as a good way of getting information without having to interact. Information sought ranged from contacts in a health support group to problems related to the use of a particular piece of software.

4. Access to expertise/experience: Half the participants said they were able to gain access to expertise and/or experience of others by lurking. Access to the expertise could come solely through the group's public forum or lead to contact outside of the group.

5. Community without posting: Half the participants indicated that a sense of community was possible while lurking. For one interviewee, a sense of community was extremely strong. This came about through a number of avenues: the interviewee's need to find community within a self help group, the stories related within the community's Web space, private postings and responses to members of the community, and the character and nature of the dialogue within the community, which engendered a sense of trust and care. The fit between the interviewee and the community was good, and in this case the outcome was a very strong sense of community, a sense that was developed without posting. Even though this interviewee has not actively lurked in the community for over a year, there is still a sense of belonging to this community.

6. Stuff in mailbox: Three of the participants wanted something in their mail box. This was described as being equivalent to receiving regular snail mail. One person indicated that they liked receiving junk mail and that receiving DL-based email filled a similar niche in their online life.

7. Connection with individuals: Lurking was described as an effective way to connect with individuals. Participants were able to determine prospective contacts from their public postings. They then went on to contact the individuals through other methods, such as side emails.

Table 4.6 describes reasons for lurking related to group characteristics. Unlike the Table 4.3: Relationships to group, this table describes the group based on characteristics intrinsic to the group, e.g., volume of messages and style of moderation. These characteristics cover a wide range of attributes that affect whether an individual lurks.

	Participant										
Group characteristics	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total
1. volume of messages			*	*	*		*			*	5
2. poor quality of messages		*	*	*	*					*	5
3. type of group (distribution vs. discussion)					*		*			*	3
4. no posting requirement (rules)	*	*					*				3
5. mechanics of interaction						*	*				2
6. style of moderation							*	*			2
7. spikes of activity					*				*		2
8. response to delurking		*								*	2
9. lack of response		*								*	2
10. delay in response								*			1
Total	1	4	2	2	4	1	5	2	1	5	27

Table 4.6: Group characteristics that were cited as reasons for lurking.

1. Volume of messages: A high volume of messages was cited by half the participants as a reason to lurk. In general, the higher the volume, the more burdensome the group became. This made following

conversations more difficult and was often related to a lowered quality of message. For the most part, groups with a high volume of messages were ones where lurking was both easier to do and deemed an acceptable form of participation.

2. Quality of messages: Poor quality of messages was mentioned by half the participants as an inducement to lurk. Quality meant different things to different participants, e.g., off topic posts, poor use of subject headings, and low information content. Communities cited as largely information interchange communities (e.g., software application help groups) were frequently left because they no longer supplied information in sufficient quantity or quality. This was largely a result of communities repeating topics and the interviewees becoming more expert in their knowledge.

3. Type of group: When interviewed, participants discussed a broad range of groups. Several participants described email-based lists which were for distribution only. In this type of list, lurking is the only way members can participate.

4. No posting requirement: Some groups have rules indicating that some level of public posting is required. If there were no specific rules, some of the participants felt more comfortable lurking. In at least one case, a participant would not belong to a DL in which participation was required.

5. Mechanics of interaction: Two issues were brought up by participants. The first concerned the use of digests to distribute messages. This made it harder to respond to a given message and thus increased the likelihood of lurking. The second concerned the operation of a BBS system in which it was difficult to follow message threads. In this particular BBS, messages could be followed down a thread but not up.

6. Style of moderation: Moderation comes in many forms. In this case, moderation causing delays (of up to 24 hours) and moderation which was heavy-handed and editorial in nature were regarded as reasons to lurk.

7. Spikes of activity: Several groups were described as having periods of rest followed by periods of high activity. During the rest periods, everyone in the group lurked. This spiky activity was considered normal for some groups and the corresponding periodic lurking was deemed to be normal in these groups.

8. Response to delurking: For several participants, how delurkers were responded to was an important indication of how receptive the group was to new members. Unresponsive or ill-mannered responses were an indication that lurking was the best approach to participating in the group.

9. Lack of response: This refers to whether the group responds to public posts. Groups where posts are not responded to were considered less inviting, and thus provided a greater reason to lurk.

10. Delay in response: Similar to the item above, groups in which there was a delay in response suggested that the group was both less active and thus more lurkable, or that the group was less capable of responding to members. In either case it was seen as a reason to lurk.

Table 4.7 outlines two reasons for lurking that are based on an individual's stage of membership. Participants articulated two periods in the life of their membership in which lurking was likely to take place. The first is when they are getting to know the group and the second is during the time when they are leaving the group.

Stage of membership	Participant										
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total
1. leaving the group		*	*	*			*		*	*	6
2. getting to know group	*	*		*			*			*	5
Total	1	2	1	2			2		1	2	11

Table 4.7: Stages of membership cited as reasons for lurking.

1. Leaving the group: For many of the participants, leaving a group was a time in which their participation was reduced. While leaving a group they read few if any messages and did no posting. Lurking while leaving was mentioned by 6 of the 10 participants.

2. Getting to know the group: Knowing the group and the individuals in the group was so important to the interviewees that many were able to describe their tactics in detail. These included looking at previous posts by an author (using archives or other means), examining email addresses for personal or corporate information, following threads to understand the nature of the discussion and participants, and using signatures and related Web sites to find out more about posters.

Table 4.8 shows the external constraints that were mentioned as reasons for their lurking. These constraints relate to time and work issues.

External constraints	Participant										
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total
1. work related	*	*	*	*	*		*	*	*		8
2. time		*	*	*	*		*				5
Total	1	2	2	2	2		2	1	1		13

Table 4.8: External constraints cited as reasons for lurking.

1. Work related: A number of interviewees listed work related reasons for lurking. In one case, the interviewee was employed to lurk. For several others, they were expected by their employers to belong to work related groups. There was no expectation by their employers that they publicly post. Many of the work-related groups were joined because the interviewees thought the groups would be of value in their work. However, given the demands of their jobs, the time and effort required to post limited their involvement within the groups to lurking.

2. Time: Half of the interviewees cited a lack of time as a reason to lurk. Being a member of any group requires time and when these participants had less of it, they lurked rather than posted. There was general agreement that posting in public spaces took more time than lurking. This was not just a case where posting required more effort, but one in which the poster was entering into a dialogue in which more posting would be required.

This ends the results section related to the *primary* question *P1: Why do lurkers lurk?* A discussion of these results can be found at the start of the next chapter. The next section presents the results to the *primary* question, *P2: What do lurkers do?*

4.4 P2: What do lurkers do?

Despite the small sample size, lurking was described as a complex set of activities (and rationales) situated in a rich space of possibilities. To make sense of the various activities, the interviewees' responses have been organized into the following categories:

- posting
- getting to know the group
- managing messages
- reading
- selecting and filtering
- finding and browsing

Each of these are discussed below. It should be noted that this is not an orthogonal set of categories and that there is interplay and overlap between them.

Posting: 8 of 10 interviewees who joined a group with the intention of lurking, lurked without ever posting. If, as is the case for several of the interviewees, they decided to lurk after having been active, they would then post only occasionally. Two interviewees did not post in any group, citing a number of reasons that included, work-related privacy and fear of receiving spam. Two interviewees said that it was easier to post a query than to respond to one as it took more expertise to respond to a query. This suggests that one can be a non-lurker when it comes to requests, and a lurker in providing a public response. Only two interviewees said they did not initiate side posts in the form of one-on-one emails to other group members.

Getting to know the group: Coming to understand the group and its members was explicitly mentioned by 7 of the 10 interviewees. It appears to be very important and interviewees were willing to expend considerable effort in doing so. Four members made specific mention of attempting to gain a general level understanding of the group and did so through FAQs, rules, and/or a related Web site.

The ongoing stream of messages and also archives were mentioned by interviewees as a good means of finding out about the tone and the rules of the group. Subject headers were used to determine the value of the list. One interviewee mentioned that they also used cross-posts to assess a group.

Knowing about individual members of a group was invaluable in evaluating the worth of a group. Knowledge about individual members came from many sources; the following were cited: name and aliases of members, quality and quantity of posts, signatures and URLs of posters, posters' email addresses, and time of day posts were sent.

Managing messages: Each of the interviewees was able to describe a method for dealing with and managing messages. Nobody read every posting, and depending on the experience within the group interviewees might not read any of the postings. Subject headings were used as a primary means of determining what to read, and the poster's name was used, if at all, as a secondary guide. Several users deleted or ignored whole threads based on a heading, well aware that heading information was often a poor indicator of content. One interviewee said this was a reasonable strategy because information tended to get recycled over a period of time.

Because the email-client collects messages from many different sources, e.g., friends, business associates, spam, DL-based email, etc, management of messages is an important task. There are a number of levels to this task, and the one that will be discussed here, is what to do with the messages. One interviewee described his management process as one of deleting almost all messages. This allowed him to maintain an inbox which had few messages and was uncluttered with extraneous messages. It also meant that he regularly went through the deletion process, usually once or twice a day. Another interviewee, interested in

keeping messages for future reference and for ease of reading, printed out all messages from a specific DL in digest format. In a somewhat different frame of mind, an interviewee, knowing that DLs recycle ideas periodically, deleted unread threads when short of time. Each of these interviewees was attempting to get the most out the groups in a way they were most comfortable.

Related to managing messages is the format in which the messages are received. DLs can send out messages as individual emails, or clumped together in digests. The theory is that it is easier to deal with a digest than many separate emails. Four of the 10 interviewees received all of their DL messages as individual emails. Four received some lists in digest form and others as individual emails, and two interviewees did not specify. Those receiving digests were committed to lurking in those groups. Two of these interviewees commented on how difficult it was to follow threads in digest format, which points out a possible tradeoff for DL managers. When setting a DL's default message format, is interactivity (in the form of following threads) more or less important than the convenience of receiving a digest?

Reading: Reading all messages was not common amongst this group of lurkers. Only one read all messages and this was because it was part of a job. Many of the interviewees indicated that they read all messages during the initial period of membership. After the initial break in period, the interviewees were more discriminating in what they read. As previously mentioned, one interviewee read and deleted all email once or twice a day. Another did it only twice a week. In the case of the daily read and deleter, reading the messages on the day they were received was important. For the twice-a-week read and deleter, the type of group (in this case a music-fan group) meant it was less important to receive the messages on the day they were sent. Whether a group is read as a newsgroup or received through a DL appears to make a difference as to how the messages are dealt with. In the case of email coming into an inbox, it appears to be very important that the messages are dealt with in some manner. In the case, where some external tool/representation tracks messages, as in newsgroups and BBSs, the act of managing messages is less imperative.

Selecting: The process of selecting messages could be considered part of the reading process, as it determines what to read. However, selecting and reading have been separated because selection represents a large and diverse set of activities. Subject headings were cited as the primary means of determining what to read. Receivers of digests said they used the Table of Contents to review the contents much in the same way subject headings were used for single email. Not having a Table of Contents made the digests more difficult to use.

The sender/author was used, if at all, as a secondary guide. A number of interviewees said that they came to recognize that certain authors tended to flame (or have other specific behaviours). Depending on the interviewees persuasion, this identification could be used to either avoid the message or read the message (often for entertainment). When messages were long, several interviewees indicated they would read the first paragraph of the message before either continuing to read on or not.

Threads were used as a means of determining value of a group of messages. If the topic of the thread was of no interest, then it was either ignored or deleted. If the topic was of interest, one interviewee said he would pick a message part way through the thread. He would then read this, determine its value, and then either reject the thread or read the thread. Several users deleted or ignored whole threads based on the heading alone, well aware that heading information was often a poor indicator of content. One participant said this was a reasonable strategy because information tended to get recycled over a period of time. Other overarching strategies were undoubtedly at play, including evaluating the worth of an email given the amount of time available.

Finding and browsing: Belonging to a group was more than just following and initiating conversation. Activities associated with belonging to a group also included browsing and finding information, expertise, and support, and in at least two cases, a sense of community was also found. Message archives were both searched and browsed. Supporting Web sites were also used. Two interviewees mentioned reading stories found on Web sites to improve their understanding of the group.

Summary

This chapter has presented the results of the interviews. Of the 41 groups mentioned during the interviews, 30 were groups where the interviewees lurked. DLs were the most frequently mentioned type of group interviewees lurked in 21 of 25 mentioned DLs. The initial assumption that the interviewees would be lurkers was sound. It also appears that DLs are a good group type in which to study lurking. The next chapter describes a study in which DLs were logged in an attempt to understand how many lurkers are out there.

These results have richly illuminated the two *primary* questions. In the case of *P1: Why do lurkers lurk?*, interviewees provided a rich array of 117 reasons. These reasons were summarized and discussed under 7 different categories. The reasons and categories will become the basis for several models of lurking in the next chapter. The results related to *P2: What do lurkers do?* indicate that lurking is more than reading messages. Lurkers participate through a wide range of non-posting activities, employing strategies and tactics in what appears to be an idiosyncratic fashion.

An in-depth discussion of the results follows in the next chapter. This discussion will be organized using both the *primary* and *related* questions developed in Chapter 2. The third *primary* question, *P3: How many lurkers are there?*, is examined Chapter 6.

Chapter 5: Discussion of interview results

Overview

- This chapter begins with a general discussion of the results.
- Results are discussed related to *P1: Why do lurkers lurk?* and two models of lurking are put forward.
- Four *related* questions are addressed:
  - R1a: What motivates lurkers?
  - R1b: What role does lurking play in learning about the group?
  - R1c: How does persistent conversation affect lurking?
  - R1d: How do individual and group character differences affect lurking?
- Results are discussed as they relate to *P2: What do lurkers do?* Five different strategies for lurking are described.
- Results are discussed as they relate to *R2a: What are the constraints on lurkers' activities?*
- Other results are examined.
- The chapter is summarized.

In reviewing the literature in Chapter 2, a number of *primary* and *related* questions about lurkers and lurking were established. The current chapter uses these questions to structure the discussion of the interview results. In the course of addressing the *primary* and *related* questions, three models for lurking are described: the *filter*, *gratification*, and *persistence* models. In addition to the questions and models, two other areas are discussed: how lurkers view themselves and lurkers' sense of community.

5.1 General discussion

The initial assumption that lurkers could be found by polling a physical community proved to be a good one. All interviewees were experienced lurkers, but not necessarily all of the time, nor in all communities (see Table 4.1). In the next chapter, further work on the demographics should provide a better understanding of just how widespread lurking is on a community-by-community basis.

Of the groups described by participants, 30 of 41 were ones in which the participants lurked. These findings support suggestions that lurking is a common activity in online groups. 25 of the 41 groups listed were DLs (Table 4.1). Of this group, 21 were described as DLs in which the interviewees lurked. Finding that DLs are widely participated in and that participation is largely in the form of lurking was one of the deciding points in favour of using DLs in a demographic study (as described in Chapter 6).

Reasons for lurking were varied, with participants citing 117 reasons (see Tables 4.1-4.8). What lurkers did while lurking was equally varied (as described in Section 4.4). The most obvious conclusion to be drawn from the interviews is that lurking cannot be characterized by the single behaviour of not posting. Lurking should be viewed as a complex set of actions, rationales and contexts, i.e., situated action. Suggesting that lurkers are free-riders as Kollock and Smith (1996) describe them is a gross simplification and appears to be untrue for these interviewees. If there is value in calling lurkers, free-riders, then it as a straw man. Others including this author will challenge this position. The limiting view of lurkers offered by the free-rider designation helped spur on the research described in this thesis.

The interviewees in this first study painted a very broad picture of why they lurked. It would be convenient for designers, community builders, etc., if specific reasons for lurking could be cited as more important than others. That does not appear to be the case with this group of interviewees. Each individual appeared to be guided by their own reasons, needs, and stage of membership. However, by examining their stated

reasons for lurking, it is possible to get a sense of the issues involved. The next section examines why lurkers lurk in detail.

5.2 P1: Why do lurkers lurk?

There is no single answer to why lurkers lurk. This is amply demonstrated by the many reasons voiced by the interviewees in this study. To understand why lurkers lurk, several models of why lurkers lurk are put forward in this chapter. This section describes two models of lurking, the *filter* model and a *gratification* model. A third model, the *persistence* model, is introduced later in the chapter.

***Filter* model:** In the previous chapter, interviewees reasons for lurking were aggregated around 7 categories. These categories are summarized in Table 5.1.

Table	Category (reasons for lurking)	Description
	member's character	
4.2	• Personal	Reasons for lurking which were personal in nature. The reasons generally reflect a desire to reduce public exposure.
4.3	• Relationship to group	Describes how individuals see themselves in relation to the group, either socially or from an informational perspective. The relationship to group either inhibits their public participation or reduces the need to participate.
4.4	• Intention from outset	These are reasons why participants lurked from the outset. Seven of the ten participants intended to lurk from the outset and gave a variety of reason for doing so.
4.5	• Wants/needs met by lurking	These are reasons given for lurking related to meeting their wants or needs. For all of the participants, lurking provided a means of satisfying their wants or needs. This is not to imply that lurking satisfied all their needs, only that it was capable of supplying a variety of wants/needs.
4.6	**group characteristics**	Describes reasons for lurking related to group characteristics. Unlike the Table 4.3: *Relationships to group*, this table describes the group based on characteristics intrinsic to the group, e.g., volume of messages and style of moderation. These characteristics cover a wide range of attributes that affect whether an individual lurks.
4.7	**stage of membership**	Interviewees articulated two specific periods in the life of their membership in which lurking was likely to take place. The first is when they are getting to know the group and the second is during the time when they are leaving the group.
4.8	**external constraints**	These are the external constraints mentioned as reasons for lurking. These constraints relate to time and work issues.

Table 5.1: Summary of categories used to collect and describe reasons for lurking.

The bulleted categories in Table 5.1 are related to the member's character. These items are what lurkers bring with them to their participation in a group. In contrast, the category following the bullets, *group characteristics*, reflects the character of each group. These characteristics have an impact on participation and in some cases give cause to lurk. The *stage of membership* category represents the ebb and flow of public and private participation as a function of the stage of membership. The last category in Table 5.1 represents constraints on participation that are generally outside of the control of the member, but have an effect on how the member participates in the group. Each of these categories represents filter or barriers to posting. The *filter* model shown in Figure 5.1. depicts the reasons for lurking as a series of filters to public posting.

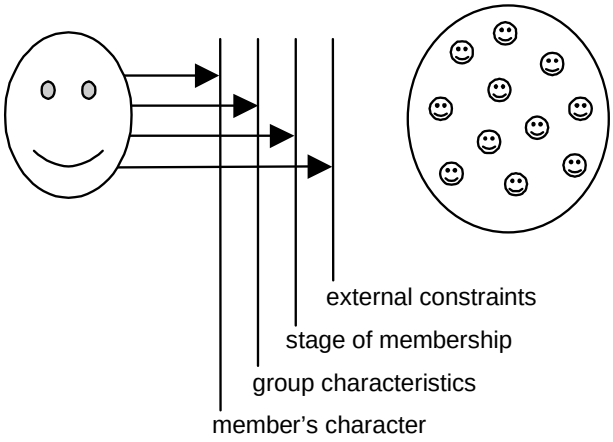

Figure 5.1: The filter model and its four filters to public posting in DLs.

Similarly, these same reasons for lurking may affect the lurkers' participation in the group, e.g., participation by reading group messages (see Figure 5.2). For example, not having enough time or nearing the end of a relationship with the group will reduce or eliminate the reading of group messages. The reasons cited for lurking can also have a positive effect on participation outside of posting. For example, in the early stages of being a member, new members are likely to read many if not all of the messages.

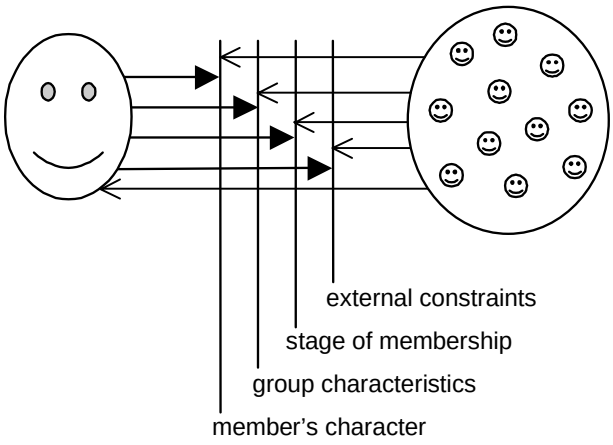

Figure 5.2: The filter model showing the effect of the filters in both directions.

***Gratification* model:** In contrast to the filter model, which views lurking in terms of constraints, the *gratification* model depicts lurking as the best means to an end. Table 5.2 lists the top reasons for lurking. In order to be included in this list, each reason was stated by at least half of the 10 interviewees (the number following each entry indicates the number of interviewees who stated this reason):

Category (of reasons)	Reasons for lurking
1. Lurking provided	conversations/stories (8)
	entertainment (7)
	access to expertise/experience (6)
	information without interaction (5)
	community (5)
2. Lurking took place because	they wanted to be anonymous, and preserve their privacy and safety (8)
	of work constraints (8)
	the volume of messages was too low or too high (5)
	of the poor quality of messages (5)
	of shyness (5)
	of limited time (5)
3. Lurking took place while	leaving the group (6)
	getting to know the group (5)

Table 5.2: Top reasons for lurking grouped in three categories.

Table 5.2 covers three broad areas. For the discussion of the *gratification* model, the third category should be ignored. The *gratification* model is based on the interaction between the first and second categories in Table 5.2. The first category indicates that lurking can provide objects or experiences. The second category describes why lurking took place in terms of either an intrinsic or extrinsic constraint (with respect to the individual). The interlinking between the first two categories is the basis for the *gratification* model. For example, an individual looking for information may wish to remain anonymous. In this situation, lurking is a means of getting their needs met. The *gratification* model is further explored in the first of the *related* questions, *R1a*.

R1a: What motivates lurkers?

The *gratification* model can be used to interpret and integrate the first and second sets of reasons in Table 5.2. Within the *gratification* model are two types of needs which lurkers sought to satisfy: personal and information needs. Interviewees described specific needs and joined groups in order to satisfy their personal and information needs. These needs varied between interviewees and depended on the context, e.g., type of group. How they satisfied their goals was also context dependent. In many instances it was possible to satisfy their needs without posting, i.e., through lurking. This gratification through lurking was not a simple process of reading every posting, but a complex, idiosyncratic process influenced by the individual's goals, experience, and the specific group in which they were involved.

For example, one interviewee belonged to a broad range of DLs, having joined them for both personal and business reasons. While the motivations for joining each list was different (e.g., want to know vs. need to know), participation in the DLs was for the most part limited to lurking. Lurking was comfortable and

enabled him to attain his goals given the nature of the DLs, each DL having high volumes of quality postings representing both depth and breadth of knowledge. In neither group was the interviewee motivated to post for information. Instead, he took a more general wait-and-see approach.

Satisfy personal needs: When DLs were joined for personal reasons there was a correspondingly strong motivation to get as much out of the DL as possible. Entertainment was a common theme and took a variety of forms. Just as some people enjoy receiving snail mail, several participants enjoyed receiving email, indicating they liked having new email in their inbox. This gave them a sense of connection and also something to do in their free time. Others mentioned being attracted to controversy and debate, including watching flaming from the sidelines. Curiosity and learning were high on many peoples' list of reasons for joining and lurking in a DL. Humour was also appreciated.

Others joined DLs with many of the same members as their non-electronic based organizations. In their opinion, this complemented and strengthened relationships. DLs also provided a convenient way to track events and announcements. One participant, who belonged to such a DL, read all messages and deleted all but the announcements for physical meetings.

Some participants are attracted to health-support DLs as a source of empathy (Preece, 1998). For at least one participant, empathy was strongly felt while lurking. DLs can also act as a mechanism for putting people in contact with one another through more private channels. For example, peers, expertise, and finding people beyond a local geographic community were described as reasons for joining a DL. Topics of specific interest to participants also drew them into joining DLs. Participants often described members of DLs as interested and focused. Relationships developed out of belonging to the DL, although no long-lasting friendships were reported, as found elsewhere (Parks & Floyd, 1996). Several participants indicated they developed a sense of community through lurking.

Members of DLs have a variety of personal needs to satisfy. These are far ranging and a number of different approaches could be taken to improve and ensure they are satisfied. These include:

- providing profiles of members (to enable contact between individuals)
- suggesting related DLs and organizations, indicating attributes and differences
- providing sets of personal stories in health-support communities.

Obtaining the above types of information and keeping it current may be more of a challenge than making it available in a usable fashion. Profile information may not be necessary for all DLs, and unless there is a proven need, may require more effort than it is worth, i.e., to collect, maintain and ensure against its misuse.

At the operational level, a means of identifying specific types of messages, e.g., announcements, moderator comments, obvious flames, would aid the lurker in sorting and using the messages more effectively. A number of DLs already employ subject header prefixes for identifying message types. For this to work, members must comply with the conventions or have a moderator determine each message's type. Knowing the conventions requires an educated poster.

Satisfy information needs: Satisfying information needs was important to the participants. In some cases, information was more important than interaction. In addition to messages, having information in the form of archives was useful to several users, especially if it was readily searchable. In a more passive way, the turnover of information through members' dialogue was also informative. In this way, participants were able to identify experts and if need be, seek expertise directly from these individuals.

Participants sought three types of information: factual information (e.g., job postings, and solutions to technical problems); different viewpoints arising from different levels of expertise; and access to personal experiences of others. Participants also mentioned breadth and depth of expertise as being important, as was finding "authentic" information based on an individual or group experience. Timely information was also considered quite important both in the sense of it being current, and that it meet the participants' immediate needs. Getting information from people living in the Middle East during the Gulf War was given as an example of timely information.

Professional needs, such as keeping abreast of conferences and work being done by peers and colleagues, were cited. Understanding who is doing what and where appears to be an important part of staying abreast of a professional community, particularly an academic one.

Artifacts and mechanisms for satisfying information needs must be better understood and their UIs improved. DL archives should be considered as information resources and their UIs should be designed to exploit this. Individuals within a DL act as living information sources; identifying expertise within a DL and making this identification known to members would aid information seeking. Message typing would be valuable for some types of information seeking, e.g., to identify profession-related announcements.

In addition to the *gratification* model, interviewees brought up two other areas related to why they lurked: ensuring privacy and safety, and reducing noise and exposure (see the third category in Table 5.2). There were two other reason to lurk: to ensure privacy and safety, and to reduce noise and exposure.

Ensuring privacy and safety

Participants were generally aware that DLs have a life of their own, and that the combination of persistence and later uncontrolled access means that there is no such thing as privacy. This inhibited their posting of personal information, and in one case, a participant's employer prohibited posting. Privacy is a concern not only at the time of the posting, but also as a long term consideration due to the persistence of DL artifacts.

Members and potential members of a DL should have a clear understanding of the implications of posting, i.e., loss of privacy. Part of that understanding lies in knowing whether the DL is publicly archived, whether there are membership criteria that have to be met in order to join the DL, and whether a list of members is readily available. At this time the majority of DLs do not provide membership lists.

Safety is also a concern for some lurkers. Participants who had concerns about safety expressed it at two different levels. The first relates to a fear of violence, i.e., that someone or some agency can use posted information (or mere membership in a DL) to find someone or something about someone. The second relates to the fact that if you don't post you can't offend, and therefore will not become a target of flaming. While the safety issue is different from privacy, the design implications are similar.

One option for ensuring privacy and safety is the use of anonymous email hosting services such as hotmail.com. These services provide mechanisms for anonymously posting and receiving messages. There is a conundrum; participants were interested in maintaining their own privacy yet wanted to know more about other members. For example, a poster's address and signature were mentioned as a means of understanding the poster, and one participant wanted to find DL members of a similar age and gender.

Reducing noise and exposure

Most participants realized that DLs and other online forums are regularly pilfered for email addresses, which are then sold or used directly to spam. Not one participant said they look forward to receiving spams. Spammers can obtain messages directly from the messages themselves or by querying the DL server for a list of members.

As a first level of defense members' addresses should be made difficult to access. Owners of DLs can easily restrict access to the DL membership list. Similarly, DL server software can be set up to prevent the distribution of email from non-members. Some DL members take their protection one step further and provide incorrect return addresses on their email. While this may foil spamming, it makes legitimate communication difficult, e.g., to get the correct address takes more effort when side posting.

DLs allow emotional detachment as the audience and thus the lurkers are for the most part not identifiable. As one participant expressed it, when you lurk, you can have curiosity without exposure. In contrast, several participants indicated that it is much more difficult to lurk in chat rooms than DLs as chat rooms are synchronous environments where participants are normally visible and thus approachable. For some participants, the practice of lurking makes leaving a DL easier in that there is less of a commitment to a DL if you don't post. For some individuals, their notoriety makes posting problematic, e.g., few government officials post to public DLs.

Some DLs discourage lurking, at least at the outset, suggesting in their introductory message that newcomers should provide a description of themselves and post it to the DL. Other DLs specifically state that posting is not required. In either case being aware of the rules of the DL is an important part of participation. Few of the participants in this study indicated that they read the rules or guidelines.

R1b: What role does lurking play in learning about the group?

As can be seen in the third group of Table 5.1, half of the interviewees lurk on joining a group. They generally take a period of time to understand the group. This period is taken in order to evaluate the group for its fit/value to them, and also to come up to speed on the individuals in the group, dialogue styles, the language of the group, and its rules (implicit and explicit). For these interviewees, lurking was a preferred method of doing this. For many, it reduces the risk of making a faux pas or being rejected.

Several participants mentioned watching how others were treated when they posted for the first time (i.e., delurked). Several aspects of the delurking process stood out for them. One is how the delurker carried out the delurking action, i.e., the presentation of self to the group for the first time. The other is how they were received by the group. It is interesting to note that lurkers evaluated the welcoming quality of the group based on the public response. They knew little if anything of the private responses engendered by the delurking. This points out how important these public spaces are to both lurker and poster alike, especially in the initial phase of joining when the group is being evaluated for its fit to the participant. Understanding whether there is a correlation between delurking and lurking rates will be examined in the next chapter.

Most participants described the process of understanding the DL as a period of intense reading of most, if not all, posts. This occurred whether the posts were available as separate emails, digests, or archives. In several cases, reading of current posts was supplemented by searching and reading archives. During this period, which ranged from days to months, participants worked at identifying the topic or topics of the DL and determining whether this was a good fit for their needs.

New DL members are inquisitive and DL owners need to take advantage of this. The following are some of the information types that could prove valuable:

- terminology dictionary
- rules, if any
- selected highlights from the archives
- selected personal stories, e.g., in health support DLs
- description of moderation (and moderators)
- topic lists
- message rate
- number of active posters
- number of members.

While some of the above are provided in the subscription or welcome message, messages of this type were frequently unread by the participants. Key to the success of this type of information is making access obvious, timely, and ubiquitous. Many of the above list items could be kept in a DL-related Web site. A link to the Web site appended to each DL message could provide access to the site. Unfortunately, having a related Web site and linking back to it is not widely practiced.

Contrasting lurking during the initial joining period is lurking when leaving a group. Interviewees were asked to describe both current and past groups. As a result, a glimpse into why they left groups emerged. Six of the ten participants indicated that they lurked in the process of leaving a group. Many indicated that a lack of time was an important element in their leaving a group. However, groups cited as largely information interchange houses (e.g., software application help groups) were frequently left because they no longer supplied information in sufficient quantity or quality. This was largely a result of groups repeating topics, and the interviewees becoming more expert in their knowledge. Others left because they were less in need of what the group could offer. For example, one of the participants belonged to a health

support group. Once his health was no longer an issue, the interviewee paid less attention to the group to the point of not reading any of the posts. This points out how lurking can change depending on the stage of membership, i.e., intense reading of posts in the beginning and virtually no reading at the end. Because the participation is in the form of lurking, this change in participation goes unnoticed in the group.

R1c: How does persistent conversation affect lurking?

As outlined in Chapter 2, persistent conversation is an important aspect of online groups and in particular those groups employing asynchronous messaging tools like DLs. The participants in this study were all aware of issues surrounding persistence of email without necessarily having thought about persistence as a factor affecting their actions. (Many of the comments made during the interviews were related to other issues, and not persistence per se, e.g., time available, minimizing effort, privacy, entertainment value, and searchability.) In reviewing the interview data, persistent messages were shown to have three major effects on lurking in DLs:

- persistent messages help the lurker
- persistent messages give reason to lurk
- persistent messages are work

These can be seen in slightly greater detail in the depiction of the *persistence* model in Figure 5.3.

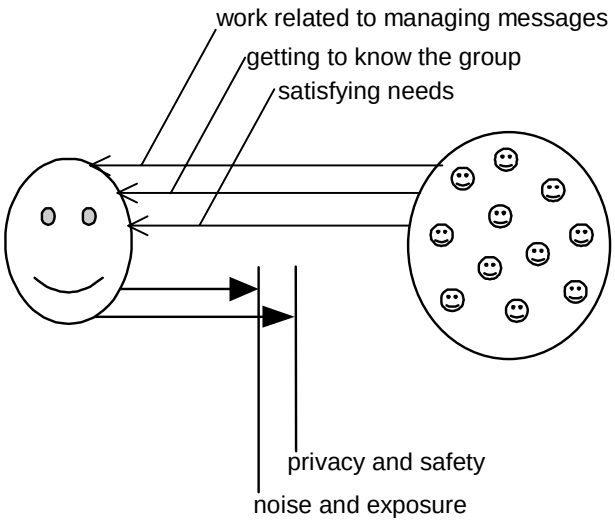

Figure 5.3: The persistence model showing the role of persistent messages in lurking.

The following describes each of these three effects in further detail. Other than the use of the negative term free-riders by Kollock and Smith (1996), the literature does not suggest that lurking is a illegitimate form of participation. As a result, it is important to understand how persistence affects lurkers in a positive way.

Persistent messages help the lurker: Persistent messages help the lurker in three ways:

- satisfy information needs
- satisfy personal needs
- get to know the group

A full description of each of the above can be found in the discussion of *related* questions, R1a and R1b.

Persistent messages give reason to lurk: The second effect suggests that persistent messages in the form of email, digests and archives are reasons why lurking occurs. Persistent messages have a life outside their intended audience for an indefinite period of time and the use of their contents is outside of the control of the participant. As a result any post is a post out of control and for anyone unwilling to let their posts have a life of their own, lurking becomes a safe option. Persistent messages give cause to lurking. Lurking ensures:

- privacy and safety
- reduces noise and exposure resulting from spam

A description of each can be found in the discussion of *related* question R1a

Persistent messages are work: The third area reflects an understanding that lurking is work, and like many other activities has management responsibilities associated with it. If the work required can be identified, then there will be a better understanding of design implications. For the DL lurker, the work is the set of actions and related time required by the lurker to deal with DL messages. For all participants, DL email was received along with other email through a single preferred email client. These email clients varied in type and configuration for each participant. As such, each participant received DL email under very different conditions. To add to the variety in both perceived and actual work, their skills with the email clients ranged from naïve to expert, and the tasks they performed ranged from simple to complex.

Participants had other priorities in their lives; DL reading/following was frequently not the most important task of the day and certainly not the one in which they wished to spend most of their time, or even a good portion of their time. In the context of their lives, lurking in a DL is one of many activities filling their day. The following strategies were employed to deal with messages.

- maximize return
- keep inbox manageable
- identify DL email amongst other email
- follow threads
- decide to read or not to read

A full description of each can be found in Section 5.3.

R1d: How do individual and group character differences affect lurking?

In the model introduced in Figure 5.1, the individual's character and the groups' characteristics are shown as filters to the public posting process. Both character sets have an effect on lurking. For example, participants' intentions with regard to public posting generally varied from DL to DL. A participant may have joined a DL with the intention of observing and never posting from the outset. If there was a mismatch between their expertise and that expressed in the DL, then this initial period of lurking was used to confirm this before unsubscribing or remaining subscribed but completely uninvolved in the DL. If there was a possibility of posting publicly, they used this period to gauge whether their posts would have value.

They also observed whether they would be able to add value, and whether the value they could contribute already existed in the DL, i.e., whether postings by others would make their contribution redundant.

While not shown in Figure 5.1, individual and group character also have an impact on the flow of messages to the user. As will be described in the next section, lurking is a methodical process and substantial effort can be invested in managing, selecting and reading messages. How this takes place is dependent on the individual and the group. For example, individuals with little or no time are less likely to read all messages from groups with high message rates. In describing acceptable message rates, several interviewees suggested that more than half a dozen messages a day is too many. It is not clear that this is the case for all users or for all types of lists. How lurking is affected by topic and message rates is examined in the next chapter.

Of the 10 interviewees in this study, at least four of them were predisposed to lurk in all of their online environments. This suggests that no matter how groups present themselves, there will be a portion of the members who will lurk. On the other hand, there appears to be a class of members who are predisposed to publicly post. It is certainly the case in many DLs that there is a constant discourse by many of the same people. Anyone who has spent time in a DL will be able to tell you approximately how many participants post daily. They may even be able to name the names of the individuals. How concentrated this group of primary posters is in terms of overall membership levels is not known. This is a question that is examined in detail in the next chapter.

Participants worked at understanding the character of the DL. They did this to increase their understanding of the DL and to become more comfortable with the possibility of submitting messages to the DL, or in several cases side posting to individuals. The term, character is used very loosely here, and includes:

- terminology or special language
- posters (players and archetypes)
- rules (implicit and explicit)
- responsibilities related to being a member of the DL (implicit and explicit)
- style(s) of interaction, e.g., confrontational, humorous, etc.
- response of members to delurkers
- style and intrusiveness of moderation
- response time to messages
- volume and quality of postings.

There are two other points related to the character of the group which have an impact on lurking. The first is related to groups in which there are periods of posts followed by mass lurking. It is as if the group collectively lurks during lulls in posting and collectively rallies into a posting frenzy on a periodic basis. From at least one interviewees' perspective, these lulls are useful as the high levels of posts take a fair amount of work to deal with and are generally quite valuable.

The other point is that large groups appear to be easy to lurk in. This is certainly true in groups that are synchronous such as MUDs and MOOs. Whether this extends to DLs, where there is little or no indication of the size of the group is not known. This too will be investigated in the next chapter.

5.3 P2: What do lurkers do?

Lurking involves many different activities. It is not just reading of posts and perhaps an occasional post as is suggested by the Jargon Dictionary. If this group of interviewees is representative of the general online population, management of messages is a very important lurking activity. The activities described by the interviewees are not passive in the sense that the reader waits for email and then responds, but involve strategies for determining what to read, delete or save. In general, activities were goal driven and somewhat idiosyncratic. Some of the idiosyncrasy may be a result of variation in tools (email-clients). For example users of non-GUI email clients such as Pine (on UNIX operating systems) tended to not use folders or

secondary mail boxes, but instead deleted messages on a regular basis. The idiosyncrasy also appears to be related to the skills of the lurker, and the goal of the lurking.

It is clear that lurkers' activities are carried out methodically and that individuals are capable of explaining not only their methods, but also the strategies they employ. For example, none of the interviewees read all messages all the time, and depending on the experience within the community they might not read any of the messages. Some interviewees were able to describe an overall set of strategies that they employed, e.g., delete all messages except for those related to announcements. Others, appeared to be much more dependent on the context of their lurking. For example, if they were short of time, they would delete whole threads, confident that the information would come up again at a later date. The following strategies were mentioned by interviewees:

1. maximize return
2. keep inbox manageable
3. identify DL email amongst other email
4. follow threads
5. decide to read or not to read

The remainder of this section describes each of these in detail.

1. Maximize return.

In general the participants were interested in getting the most out of the time they had for lurking. Even if they lurked to entertain themselves, they still wished to do this as efficiently as possible. This typically meant spending less rather than more time with the DL(s). They used a number of methods to do this. If they belonged to more than one DL, they limited themselves to the number of DLs they could handle. It was clear that too many DLs meant that the value of one or more of the DLs would be reduced.

While many of the DLs described by participants had 20-30 messages/day, participants were generally happier with fewer messages. Factors affecting the amount of time required to lurk on a DL include the quality and size of the messages, the motivation in belonging to the DL, the volume and type of email received from other sources, and the time available. In the examination of a number of introductory messages provided by the DL and DL related Web sites, none mentioned how many messages a subscriber might expect.

The asynchronous and persistent nature of DLs means that lurkers can go back through archived messages at any time and either search for particular information or browse the messages. For some lurkers this is an efficient way of finding pertinent information.

2. Keep inbox manageable

Manageable meant different things to different participants, but was often related to comfort. For several participants comfort came from keeping their inbox small, i.e., able to see all retained messages at once. The process of picking through the messages was an important part of their management process. Understanding how inboxes are used is critical to developing design solutions.

The use of filters to sort messages into secondary mail boxes was not commonly used among participants. A number of reasons were stated: not trusting the effectiveness of the filters, potential burying of important email, and no knowledge of filtering tools or the process of creating effective filters. Filtering mechanisms should be examined with an aim to making them verifiable, trustworthy, and simpler to learn and use.

3. Identify DL email amongst other email

Differentiating one DL's messages from another, and those in turn from non-DL email was an effort for participants. Recognizing this as a problem, some DLs use an identifying prefix in the subject header to indicate that a message is from a particular DL, e.g., the MORE cycle DL prefixes all subject headers with "more:".

Identification of DL messages is an important mechanism for scanning and processing email in the inbox and elsewhere. The current ad hoc approach of using prefixes may be good enough, but could be improved upon. A related issue, although not raised by the participants, is the use of prefixes to identify different types of messages, e.g., "Q:" for question. The use of prefixes helps identify a message's origin and intent, but may also make the subject heading more difficult to read.

Existing header information is sufficiently descriptive for use in separating messages from different DLs and non-DL email. However, filtering tools remained largely unused by the participants. Whittaker and Sidner (1996) found the inbox to be an important repository for messages. Their findings suggest that the low use of filters may not reside solely in the act of filtering, but on other factors, such as the fear of losing track of important information.

4. Follow threads

A thread is a conversation of multiple messages linked via a repeatedly used subject header. Participants were able to follow threads in newsgroups and BBSs because these systems were designed with threaded conversation in mind. Participants used threading to either follow a particular discussion or determine whether a line of discussion was worth reading. This particular facility is poorly implemented or non-existent in most email clients. In addition, threading in email clients is different from that in newsgroups or BBSs. Even when messages can be sorted by subject header in an email client, the results are presented as a list of messages related by subject header. In both BBS and newsgroups, messages are related in a tree like manner, with the relationships between individual messages being apparent to the user. For this reason, email-based threading might better be called clumping.

For threading to be of value in email clients, threading must be effectively represented in the UI, e.g., threading based on subject header and date, and keeping the most active threads in the most visible position in order for the thread activity to remain observable to the user. GUI-based email clients can show threading based on the subject headers, but the results are frequently cumbersome and confusing. Alternative solutions need to be examined.

Additional problems occur when receiving DLs as digests. While digests reduce message clutter in the inbox, they eliminate thread visibility. Current email clients are unable to show threading in digests although specialized digest readers such as Digester (TECHWR-L, 1999) show promise in this area.

5. Decide to read or not to read

Determining what to read is an important activity for any lurker. Deciding whether a message was worth reading was idiosyncratic and for a given participant often differed between DLs. The following criteria were described:

- read all if participant is new to the DL
- read if the subject heading shows potential value
- read if the author is known
- read all messages in a thread if the middle message of a thread is interesting
- read messages if thread is long (i.e., quality of messages and thread is somehow related to the length of the thread)
- read messages with confusing subjects
- read or not read an obvious flame

Several participants deleted all or most messages (read or not read) as a matter of course, whereas others kept messages, either by leaving them in the inbox and relying on the read flags to indicate their status, or by manually placing them in secondary folders. The delete process was most common among users of text-based email clients.

A rich set of cues were used in deciding whether to read a message. The fact that messages are persistent and asynchronous, means that a message does not have to be read at the time of receipt. It also means that

the decision as to whether a particular message is read will often be based on other messages, e.g., other messages in the thread or the quantity of messages in the inbox.

R2a: What are the constraints on lurkers' activities?

Apart from the areas already discussed, such as the stage of the membership and the fit between member and group, constraints on lurkers activities were primarily time or work related. These constraints are largely external in nature. For example, one person was prohibited by his employer from posting and another interviewee belonged to a group because it was part of his job. In this latter case, the employer did not require or even expect the interviewee to post.

Like anything else in life which takes time, when a person is busy in one area, the time available for another area is reduced. Being involved with a community through posting can take more time. In recognition of this, several interviewees mentioned that they did not post because they knew that it would likely lead to a dialogue, which in turn would require them to continue. There is an awareness that posting has responsibilities. One interviewee indicated they were uncomfortable when someone would post to a group and a response was not forthcoming. They mentioned this was a situation in which they considered delurking or sending a private response. Whether lurking levels are related to the responsiveness of the groups is examined in more detail in the next chapter.

Several interviewees made an interesting observation. They indicated that it is much easier to ask a question than to respond to one. This was particularly true if the group was technical in nature. This would suggest that a member of a group may lurk as a respondent and post as a questioner. There is no evidence in this study to indicate this is so, but could be investigated through a combination of logging and discourse analysis.

One constraint that is typically under the control of both the DL owner and the member is the format in which the messages arrive. The interviewees who mentioned receiving DL messages as a digest mentioned they were more difficult to follow and less exhaustively read. As such, a digest could be considered a barrier to participation in all its forms, not just to posting.

5.4 Other observations

Till now, the discussion has been structured around the *primary* and *related* questions. Two other items came up during the interviews that fall outside of the questions, but are important enough to be discussed: the development of a sense of community by interviewees who lurked, and lurkers' own sense of their lurking.

Sense of community while lurking

In the process of doing the interviews, it became obvious that knowing a group and the individuals in it was very important to several of the interviewees. These interviewees were able to describe their tactics in detail. These included looking at previous posts by an author (using archives or other means), examining email addresses for personal or corporate information, following threads to understand the nature of the discussion and participants, and using signatures and related Web sites to find out more about posters.

In addition to becoming familiar with the group or groups, several interviewees developed a sense of community while lurking. This goes against the preconceived notion that you must be an active poster to be part of a community. For one interviewee, a sense of community was extremely strong. This came about through a number of avenues: the interviewee's need to find community within a self help group, the stories related within the community's Web space, private postings and responses to members of the community, and the character and nature of the dialogue within the community, which engendered a sense of trust and care. The fit between the interviewee and the community was good, and in this case the outcome was a very strong sense of community, a sense that was developed without posting. Even though

this interviewee has not actively lurked in the community for over a year, there is still a sense of belonging to this community.

Lurkers on lurking

Part of this study was to understand how interviewees viewed lurking in general and their own lurking behaviour in particular. An initial abstract for a paper (Nonnecke & Preece, 1999) was distributed to some of the interviewees. One interviewee responded with the following comment:

> Maybe it's a sign of my own mild discomfort around being a lurker, but I found it reassuring to recognize myself and my behaviour within the continuum you describe, and to see lurking treated seriously, with both acceptance and respect. As a lurker, I'm used to observing from the sidelines and participating vicariously, and it's strangely gratifying to read an article that speaks directly to that experience. It's almost like suddenly feeling part of an (until-now) invisible community of lurkers.

This interviewee was not alone in feeling there is a stigma associated with lurking, although the degree of stigmatization varied from individual to individual. Giving lurkers recognition as valid participants (beyond the current Jargon Dictionary) will benefit both lurkers and the community as a whole. Simple math indicates that if lurkers delurked, communities in their present form would become a chaotic message ground, perhaps mimicking many newsgroups in their level of disorder.

5.5 Summary

Interviewing a relatively small number of online group participants under the assumption that many would be lurkers turned out to be a fruitful approach. The interviewees provided a wealth of information on their participation in groups specifying why and how they lurked. It is now possible to discuss lurking within the framework of the *primary* and *related* questions. In addition, three models presented in this chapter can now be used to discuss lurking.

The study was intended to provide a better understanding of lurkers and lurking, and was not intended to be exhaustive or by any means the last word on the subject. Rather it was an initial probing, which could become the basis for further work. As will be seen in the next chapter, many of the questions raised in the interview study are further examined in a log-based study.

The interviews were designed to gain as broad a picture as possible. To do this, the interviewer took a practical liberty. While generally following a script, questions in the open-ended interviews focused on issues brought up by the interviewee. These new areas of interest would form a background of investigation for subsequent interviews. As a result, the interviews were less like a scripted questionnaire and had a much more fluid quality.

There are two caveats about the results which need to be understood. The first is the small sample size and the second is the nature of the sample. A larger sample may or may not provide a broader picture of lurking, but could provide evidence of the strength of the observations and conclusions. This suggests that these results could and should form the basis of larger study, one that could be based on either a very broad survey or a series of surveys to illuminate specific questions. As for the nature of the sample, it needs to be understood that by drawing participants from a well educated set of two populations (all had university degrees) has biased the results. For example, it is clear that the interviewees in this study were very reflective and that this undoubtedly affected their lurking. At some level this bias is acknowledged in the way the results have been presented. That is, the results have not been weighted for particular importance. Instead, they are presented as equally interesting. Continued research in this area will undoubtedly uncover further parts of the puzzle and will likely suggest which parts of the puzzle are more or less important. This will certainly be of interest to anyone wishing to commercially make the most of lurkers.

In the next chapter, a demography is carried out on a number of DLs. The *primary* question, *P3: How many lurkers are there?*, and several *related* questions derived from the literature will be used to structure the results and discussion. As well, questions raised in the current chapter will resurface in the next chapter.

Chapter 6: Demography results: Counting the lurkers

Overview

- The logging method used in the demographic study is described.
- An overview of the results is presented.
- Each research question is addressed in turn.
- Further discussion is presented.
- A summary is provided.

This chapter describes a quantitative study of lurkers in DLs. In Chapter 2, a number of *primary* and *related* questions were raised relating to lurkers and lurking. One *primary* questions still remains, *P3: How many lurkers are there?* and addressing this questions is the main goal of this chapter. There are other goals which can best be understood through a series of *related* questions. The first set of questions are derived from the discussion of the interview results (Chapter 5):

- Do different types of DLs differ in their lurking levels?
- Is there a relationship between lurking and the number of members in the DL?
- Is there a relationship between lurking and the traffic level of the DL?
- If posting is concentrated with a few posters, how does that affect lurking levels?
- Is lurking related to the response delurkers receive?

In addition to the above, others questions are addressed in this chapter (see the bulleted list below). The first of these examines part of the definition of the term lurker supplied by the Jargon Dictionary. That definition describes a lurker as someone "who posts occasionally or not". The second and third questions reflect the work of Whittaker et al. (1998) in which they examined newsgroups for mass interaction. In this study, measures similar to Whittaker et al.'s will be used to understand lurking. The three *related* questions are as follows:

- If lurking is defined as no posting, what happens to lurking levels when the definition is broadened to include minimal levels of posting, e.g., 1 post/month?
- Are short messages related to lower levels of lurking?
- If clumpiness (threadedness) is an indication of interaction, does it necessarily follow that increased clump size is related to decreased lurking?

Questions will again be used to structure the results, however, both the results and the discussion will be dealt with in the same section. A summary discussion follows the results sections.

6.1 Method

The work reported in this chapter is the second of two related studies. The first study described in Chapters 4 & 5 provides insight into why and how people lurk. Parts of the first study were published in two conferences (Nonnecke & Preece, 1999, Nonnecke & Preece, 2000b) and copies can be found in Appendices F & G. In this chapter the *primary* question, *P3: How many lurkers are there?* is addressed. A portion of this work has also been published (Nonnecke & Preece, 2000a) and a copy of that paper can be found in Appendix E.

The first study found that each participant lurked in at least one online group, and several lurked in all of their online groups. This finding, among others, reinforced the need to better understand lurking. A demographic survey of online discussion groups should provide a different perspective from the first study by emphasizing quantitative measures.

DLs, rather than BBSs or newsgroups, were chosen as the basis of the logging study for a number of reasons. For the results to have their greatest value, the chosen communication technology needed to be widely used. L-Soft's usage figures show very high levels of use, and of the online discussion groups mentioned by participants in the first study, 25 of the 41 were DLs accessed through email. Just as importantly, DL servers track membership through their subscription mechanism. In turn, DL membership information can be accessed by querying a DL's server. The level of lurking can be measured by tracking posted messages and identifying posters. In contrast, membership levels are unavailable for most BBSs and newsgroups.

The primary aim of this study is to understand how much lurking occurs in DLs. Building on previous work on self-help health support communities (Preece, 1998; Preece & Ghozati, 1998) it was decided to also compare lurking in health-support to software-support DLs. McMillan provides four reasons for studying health-related communities (McMillan, 1998):

> …health and health related subjects have in the past played a central role in the early financial support in many media; health related sites are the fastest growing topic areas in CMC; health-related sites are heavily used; and this area contains one of the fastest growing categories of consumer advertising.

For these reasons, health-support DLs are the focus of this investigation. For comparison purposes, software-support DLs are also included in this study.

To select DLs for the study, L-Soft's CataList catalog and DL search facility (L-Soft International, 1999a) were used to locate suitable DLs. A search on the word "support" resulted in a listing of 300 DLs and a description of each. From this listing, subscriptions were taken out on all public DLs relating to health or software-support. To increase sample size, additional subscriptions were taken out on a random selection of health (22) and software (10) support DLs. Although the additional DLs provide support for their members, neither their title nor their one-line catalog description contained the term "support". (Note: Analysis comparing these additional support DLs shows their lurking levels are not significantly different from those found through searching on "support", and as a result, they are included in this study.)

In addition to DLs related to health and software, a random set of DLs on other topics were selected for their large size (CataList displays a description of all DLs with membership greater than 1000 (L-Soft International, 1999b)). Eleven randomly selected DLs between 1000 and 2000 members were included as a basis for examining whether large DLs have a greater proportion of lurkers than smaller ones.

Messages were collected from the selected DLs over a three-month period at a rate of slightly less than 2000 messages/day. Eudora Pro was used to collect and filter email into separate mail boxes for each list, and to monitor the process on a regular basis. Using CataList, the membership size of each list was determined at the beginning and end of the collection period. Lurking levels reported in this study are based on the lower of the two membership levels recorded for the 12-week period.

At the end of 12 weeks, the DLs were examined to ensure that each DL had sent at least one post a month for the 12 weeks. Of the 135 original subscriptions, 109 DLs are included in the study. DLs were dropped from the study if they stopped sending messages for any reason, e.g., change of server, failure on the part of the researchers to reply to subscription notices, or a non-active DL. Messages from the remaining DLs were then run through a Perl script producing records containing the following fields: list name, date, time, size of message, subject heading, and sender. Messages records were imported into an SQL database. This provided an effective and flexible means for querying and analyzing the data.

A group of 10 measures was calculated from the data collected for each DL and where appropriate means were calculated for each. Table 6.1 shows the measures used in addressing each question.

	Question	Measures used to examine question
P3	How many lurkers are there?	• Lurkers with zero posts (% of members)
R3a	Does lurking in health and software-support DLs differ?	• Lurkers with zero posts for each DL set (% of members)
R3b	If lurking is defined as no posting, what happens to lurking levels when the definition is broadened to include minimal levels of posting, e.g., 1 post/month?	• Lurkers with 0 posts (% of members) • Lurkers with 1 or fewer posts (% of members) • Lurkers with 2 or fewer posts (% of members) • Lurkers with 3 or fewer posts (% of members)
R3c	Is there a relationship between lurking and the number of members in the DL?	• Lurkers with zero posts (% of members) • No. of lurkers with 0 posts • No. of members
R3d	Is there a relationship between lurking and the traffic level of the DL?	• Traffic (posts/day)
R3e	If posting is concentrated with a few posters, how does that affect lurking levels?	• Posters creating ¼ of posts (% of posters)
R3f	Are short messages related to lower levels of lurking?	• Mean message length (lines)
R3g	If clumpiness is an indication of interaction, does it necessarily follow that increased clump size is related to decreased lurking?	• Mean clump size (messages) • Number of topics (or No. of clumps)
R3h	Is there a relationship between the number of singleton posters and level of lurking?	• Singleton posters (% of posts) • Singleton posters - no response (% of posts)

Table 6.1: Questions and the measures used to address them.

The following explain the specialized terminology used in Table 6.1. This terminology will be used in discussing the relevant questions.

Posters creating ¼ of posts (% of posters) (R3e): This measure was used by Whittaker et al. (1998) in their examination of mass interaction in newsgroups. In effect, it is a measure of stardom, i.e., the concentration of posting by a few members. A low value indicates that only a few posters make up ¼ of the total number of posts , i.e., the stars of the DL create a large proportion of the total posts. A high value indicates that posting is more widely distributed, and thus fewer stars. For example, if the value is 3%, this means that 3% of the posters make ¼ of all the posts. As this number is low, it suggests that there are star posters who contribute many of the posts. If the value is 20%, then 20% of the total posters account of ¼ of

all posts. This indicates that there are few stars as the posting is nearly evenly distributed amongst all the posters.

Mean message length (R3f): For each message collected the message length was determined (less the header but including any signature). For each DL an average message length was calculated. These values were then used to calculate the mean message length for each DL set.

Clump (in R3g): DLs differ from newsgroups and BBSs in their display of threads. In newsgroups and BBSs related messages are generally displayed in threads. In email-based DLs, threads are not readily apparent except through common subject headings. Even when a common subject heading is used to organize threads, the messages appear as clumps of messages and show none of the branching associated with news readers or BBS user interfaces. If so inclined, the DL member must manually piece the threads together. For this reason, a "thread" in a DL is better described as a clump of messages.

Clumpiness (R3g): Just as threadedness is a measure of interactivity in a newsgroup or BBS (Rafaeli & Sudweeks, 1997; Whittaker et al., 1998), clumpiness is a measure of interactivity in an email-based DL. Clumpiness is the average number of messages related through a common subject heading. For example, if a DL had a clumpiness of 3, this would mean that on average, there were 3 related messages per clump. The higher the value, the greater the interactivity.

Singleton poster (R3h): This is a poster who posts once and then never again. Whittaker et al. (1998) describe these posters as "singleton posters" and found they account for 27% of posts in newsgroups. This is a rough approximation of the number of posters who delurk, send one message and never post again. Several interviewees in the first study indicated they actively observed how groups responded to members' delurking. Without doing content analysis, it is difficult to determine whether a given post is a form of delurking. However, the next best measure is a count of how many people post only once.

Singleton poster - no response (R3h): In this case the poster posts once, receives no response from the group, and does not return. As in personal relationships, ignoring someone, may be the single most effective way of putting them off further interaction. If a response is not forthcoming, this may be an incentive for lurkers to continue lurking, i.e., the group may not be welcoming.

6.2 Overview of results

In the three months messages were collected, 147,946 messages were collected from the 109 DLs. This represents over 60,000 members and 19,000 posters. As a preview of the results that will be discussed in Section 6.3, Table 6.2 presents the results for each of the *primary* and *related* questions.

	Question	Result from logging study
P3	How many lurkers are there?	Fewer than expected: still high with an average of over 55% for all DLs (when defined as 0 posts in 3 months).
R3a	Does lurking in health and software-support DLs differ?	Yes: health-support groups have lower levels of lurking (45% vs. 82%).
R3b	If lurking is defined as no posting, what happens to lurking levels when the definition is broadened to include minimal levels of posting, e.g., 1 post/month?	Lurking increases rapidly and then levels off as definition is broadened. Health-support groups maintain their lower levels of lurking (75% vs. 97% for software when lurking is defined as 3 or fewer posts/3 months).
R3c	Is there a relationship between lurking and the number of members in the DL?	Yes: smaller DLs have fewer lurkers.
R3d	Is there a relationship between lurking and the traffic level of the DL?	Yes: higher traffic means lower lurking.
R3e	If posting is concentrated with a few posters, how does that affect lurking levels?	The greater the concentration, the less the lurking.
R3f	Are short messages related to lower levels of lurking?	Yes: short messages are related to lower levels of lurking.
R3g	If clumpiness is an indication of interaction, does it necessarily follow that increased clump size is related to decreased lurking?	Yes: larger clumps are related to lower levels of lurking.
R3h	Is there a relationship between the number of singleton posters and level of lurking?	Yes: as the number of singleton posters rises (and those who do not receive a response), so does the lurking.

Table 6.2: Overview of results ordered by question.

Summary data, including frequency tables, independent sample tests and Pearson's correlations can be found in Appendices A-D. Information from these tables is used throughout the remainder of this chapter.

6.3 Specific questions: Results & discussion

The *primary* question and each of the *related* questions are discussed in detail in this section.

1. Lurking levels

P3 How many lurkers are there?
R3a Does lurking in health and software-support DLs differ?

Using information from the SQL database, mean lurking levels were calculated for the set of all DLs, and for each of the health and software DL sets (see Table 1, column 1). Lurking was defined as no posts within the 12-week collection period. The mean lurking level for all DLs is less than the 90% figures reported by Katz (1998) and Mason (1999). It should be noted that while the mean was less than 90%, 12% of the DLs had lurking levels higher than 90%.

	1. Lurking (% of membership)			2. No. of members			3. Traffic (posts/day)		
DL set	Mean	SD	SE	Mean	SD	SE	Mean	SD	SE
All N=109*	55.5	29.6	2.8	551	678.3	65.0	16.2	18.4	1.8
Health N=77*	45.5	28.7	3.3	398.4	439.9	50.1	18.4	18.4	2.1
Software N=21*	82.0	13.9	3.0	662.4	1091.2	238.1	3.1	4.7	1.0

* No. of DLs in set, SD=Standard Deviation, SE=Standard Error of Mean

Table 6.3: Lurking, no. of members and traffic for the DL sets.

The differences in mean lurking levels between the health and software-support DLs is significant. Software-support groups had almost double the number of lurkers. Figure 6.1 shows the distribution of the lurking levels for each DL type using a box and whisker display. (Note: See Sternstein (1996, p. 37) for further information on this visual representation.) Each horizontal line represents a boundary for 25% of the DLs in the sample. The thicker line is also the median for each type. Each of the central boxes contains 50% of the DLs. Two things can be seen in Figure 6.1. First, software-support DLs cover a much smaller range of values, none being below 50%. And second, more than ¾ of the health-support DLs have lurking levels that are lower than those of ¾ of the software support DLs. These distributions look very different.

Health-support is a broad umbrella under which to investigate group behaviour. As such, lurking levels may vary according to a number of other factors, e.g., list topic, illness vs. injury, or chronic vs. short term disorders. This difference in variation between the two DL types may be the result of the greater number of health-support DLs in the study, which may represent a broader cross-section of their type.

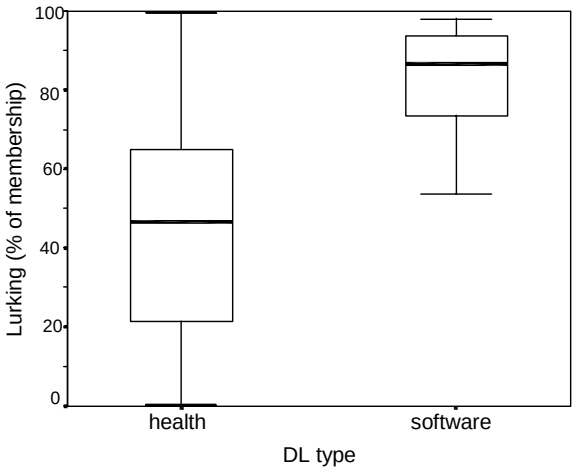

Figure 6.1: Distribution of lurking levels by quartile for each DL set.

Apart from the low mean number of lurkers in the health-support DLs, what appears most striking about these results is the large variation in lurking levels, and that on average the lurking level for all DLs is lower than the reported 90% figure (Katz, 1998; Mason, 1999).

	Question	Snapshot result
P3	How many lurkers are there?	Fewer than expected: still high with an average of over 40% for all DLs.
R3a	Does lurking in health and software-support DLs differ?	Yes: health-support groups have significantly lower levels of lurking.

2. Broadening the definition of lurking

R3b If lurking is defined as no posting, what happens to lurking levels when the definition is broadened to include minimal levels of posting, e.g., 1 post/month?

This *related* question is not derived from the previous chapter, but falls out of the Jargon Dictionary's definition of "lurker"(described in Chapter 1). That definition describes a lurker as a person who posts infrequently or not at all. Given the type of data collected in this study, it is possible to examine lurking levels at various levels or infrequent posting, e.g., from 0 to 3 posts in a 12 week period.

In Table 6.3 lurking was defined as no posts during the 12-week collection period. If lurking is examined on a sliding scale where the allowable posting level can grow, a somewhat different picture emerges. In Figure 6.2, lurking levels were calculated for a range of cumulative posts, from no posts to 3 or fewer posts for the 12-week period (i.e., 1 or fewer posts per month). As the definition broadens to include more posts in the 12-week period (towards the 3 level), lurking levels move higher. At the level of 3 or fewer posts per 12-week period, the mean lurking level for the health DLs is still lower than 90%, while the software DLs'

mean has moved above this level. Both the health and software-support DLs behave in a similar manner, and their relative offset is maintained.

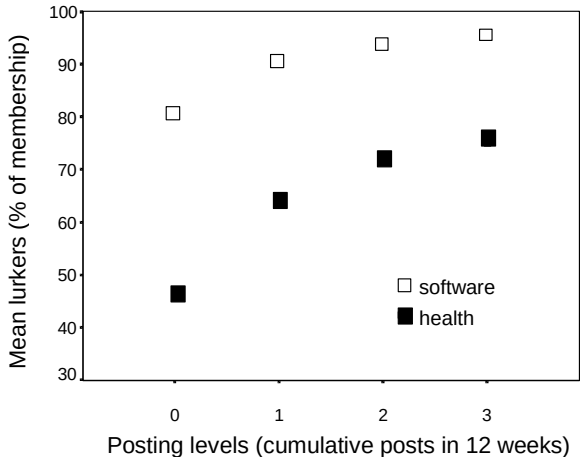

Figure 6.2: Variation of lurking levels for a range of cumulative posts.

	Question	Snapshot result
R3b	If lurking is defined as no posting, what happens to lurking levels when the definition is broadened to include minimal levels of posting, e.g., 1 post/month?	Lurking increases smoothly when broadening the definition. Health-support groups maintain their lower levels of lurking .

3. Lurking and the number of members

R3c Is there a relationship between lurking and the number of members in the DL?

In large DLs lurking may be easier. As the number of members increases, the need for any given member to participate may decline. In addition, high posting levels could create chaos and lurking in large DLs may be a practical means of reducing the number of posts and maintaining order. Several interviewees in the first study indicated they lurked when they knew that others would respond to a query. As a result, it is a reasonable expectation that large DLs should have a greater proportion of lurkers than smaller ones. As can be seen in Table 1 (column 2) health-support DLs have on average fewer members than the software-support DLs. If increasing membership size has the effect of generating more lurkers, then the difference in mean membership levels could explain why health-support DLs have lower levels of lurking.

On examining all 109 DLs in the sample, the anticipated greater incidence of lurking in larger DLs is not strongly shown. Figure 6.3 shows a strong positive non-linear relationship between the number of lurkers and the size of the DL. A linear regression also fits this data equally well (with both normalized using

ln(x+1) and ln(x), R=.76, p<.001). If this result is taken at face value, membership size does not explain the differences in lurking between the health and software DLs.

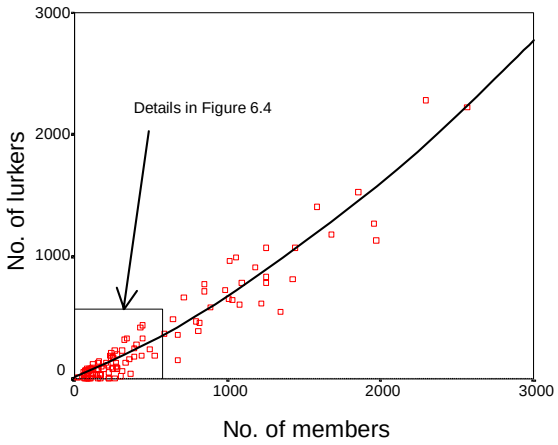

Figure 6.3: No. of lurkers vs. no. of members for each DL.

The relatively few DLs with over 500 members skews the relationship in favour of the larger DLs. Of the 98 health and software DLs, 74 of them have fewer than 500 members. Figure 6.4 is a scatter plot of these smaller DLs. The regression line in Figure 6.4 is a strong positive relationship with a slope less than that in Figure 6.3. This suggests that for DLs with fewer than 500 members, there are on average fewer lurkers than in the larger DLs. It should also be noted that the software-support DLs in Figure 6.4 are distributed in a straight line. This suggests that software-support DLs, will on average, have higher lurker levels than their equivalent sized health-support DLs.

DL members receive no direct information about the number of members in a DL. The cues that do exist are indirect, e.g., a query to the server for information, the number of different members posting, the variety of topics covered, and the traffic on the DL. It is possible that a DL of several thousand members could behave like and be indistinguishable from one with only 100 members. More work is required to understand how the size of DLs is perceived by members, and how members respond to this in their various forms of participation.

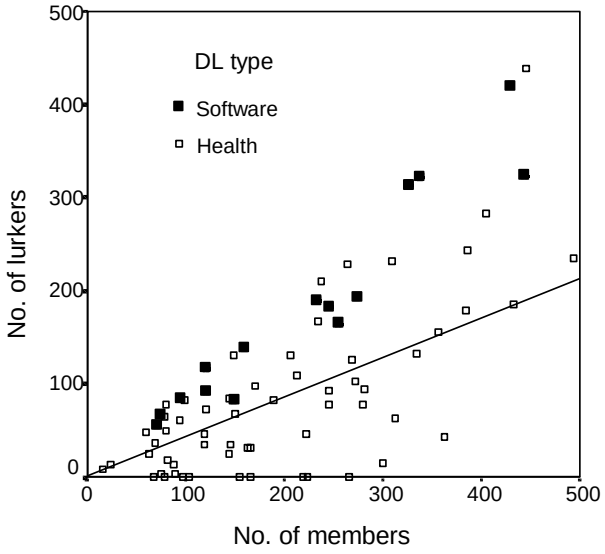

Figure 6.4: No. of lurkers vs. no. of members for each DL with less than 500 members.

	Question	Snapshot result
R3c	Is there a relationship between lurking and the number of members in the DL?	Yes: smaller DLs have fewer lurkers.

4. Lurking and DL traffic levels

R3d. Is there a relationship between lurking and the traffic level of the DL?

Two interviewees in the first study voiced an opinion on how much traffic was too much. Both said more than 6 messages a day was about the upper limit of what they wanted to receive. From the perspective of personal email management, once message rates get above a comfortable level, participating in a DL may take more effort, i.e., there are more messages to read, skim, reply to, etc. Based on the input from the first study, traffic levels were divided into four categories requiring varying levels of management effort (see Table 6.4).

It should be noted that this is a first cut at attributing management effort to various traffic levels. Other interviewees may have suggested different levels and it is likely that for some DLs, the perceived effort could be quite different. For example, a DL which might be supplying very valuable messages at high volumes (e.g., a support community at a time when the member is in crisis). Other elements that could effect management effort include type of email client, experience, demands on time, and interest.

The categorization was done prior to examining the distribution of posting rates from the current study. In the logging study, more than 50% of the DLs fall in the High category.

	Traffic level	
Management effort	messages/day	messages/week
None	< 0.14	< 1
Low	0.14-0.5	1-3
Medium	0.5-6.0	4-42
High	>6.0	> 42

Table 6.4: Traffic levels for a DL and the corresponding management efforts.

Lurking levels for all DLs were negatively correlated with traffic. (R=-.621, P<.001, traffic data was normalized using ln(x)). Figure 6.5 shows that for a given DL size, lists with the highest traffic levels generally have the lowest lurking levels. Banding by traffic level is visible, starting with the lowest traffic level (None) in the top left hand corner and progressing towards the lower lurking levels and larger DL size. This partially explains the lower levels of lurking in health-support DLs as these had the higher traffic levels (see Table 1, column 3).

Conspicuously absent are DLs in the area below the broken line, which appears to be a kind of interactive no-man's land. Why this should be the case is not known at the present time, but it could be related to the difficulty of making sense out of large DLs with high traffic volumes and large membership levels. At some point, the DL may become unusable and self-adjust through membership attrition and/or a decrease in public posting. It may be that lurking increases under conditions where having a public voice is difficult. In our initial study (Nonnecke & Preece, 1999), several participants indicated they knew other people would post opinions similar to their own and thus felt no need to post. When traffic is high, there is a sense that adding messages to the list only increases the traffic without improving the quality. For them, lurking was a way of reducing the noise on the list, a civic duty so to speak. It would be interesting to examine DLs that fall near or below the broken line, and determine whether they transform in any way, e.g., split, have high membership turnover, etc.

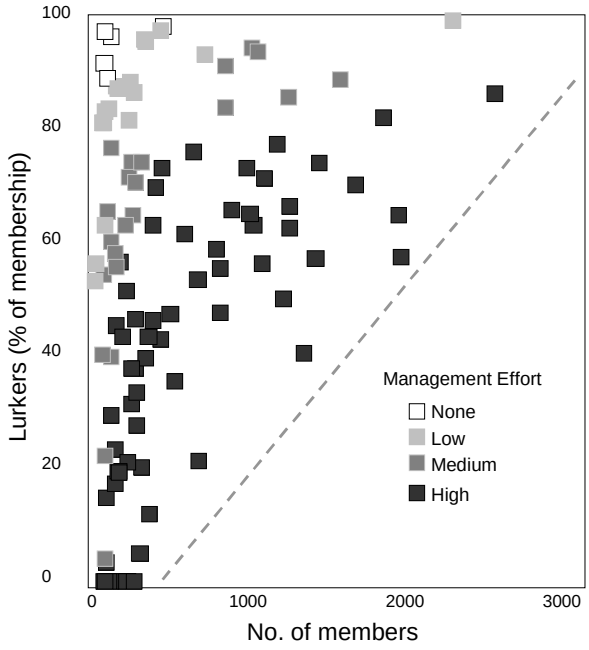

Figure 6.5: Lurker levels as related to management effort and number of members for each DL.

Below the 500-member level, health-support DLs appear evenly distributed with respect to the number of lurkers and thus lurking levels (see Figure 6.4). For these smaller, more personal-sized groups, the size of the DL may be less of an indicator of lurking level and some other factors may be at work. For example, for DLs with fewer than 500 members, traffic levels appear to be a good predictor of lurking levels (see Figure 6.5). What drives the combination of low lurking levels and high traffic is still unclear, but may be related to the topic of the DL, motivation of members, and style of interaction (e.g., empathy vs. information exchange).

The DLs with high traffic levels are an interesting group (see Figure 6.6). The 11 DLs with average traffic levels over 40 messages/day had a low average lurking level of 44%. Four of the DLs were from the Large set of DLs and 7 were health-support. The median membership size for this group was high, at 1220. However, three of these high traffic DLs had fewer than 500 members (and all were heath-support DLs). For the DLs in this high traffic range, lurking levels appear to be randomly distributed across membership size. As a result, high traffic levels don't appear to be a very good indicator of group size. It is possible that group size becomes immaterial to public participation when it isn't directly knowable, as it is in a typical DL. However, as group size increases, so does traffic. (R=.410, P<.001). It would not be surprising for members of DLs to equate high traffic levels with large memberships.

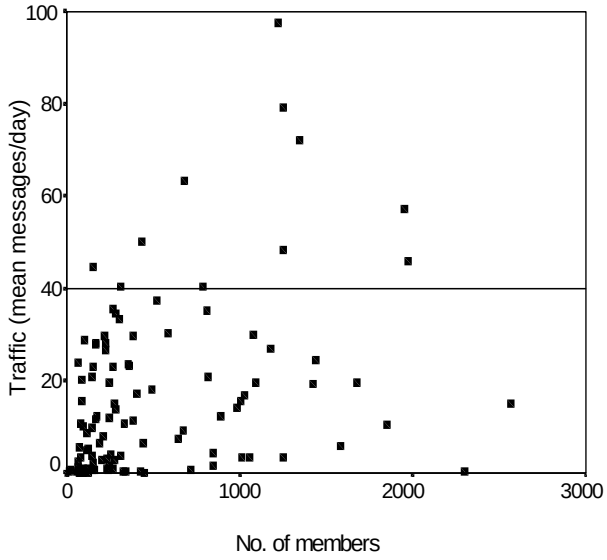

Figure 6.6: Traffic levels vs. no. of members.

	Question	Snapshot result
R3d	Is there a relationship between lurking and the traffic level of the DL?	Yes: higher traffic means lower lurking.

5. Lurking and stardom

R3e If posting is done by a small percentage of posters, how does that affect lurking levels?

This is a question of understanding the demographics of those who post and then determining whether there is any evidence that lurking levels are somehow related to the distribution of posts. For example, will a DL with a small number of heavy posters have fewer or greater number of lurkers than a DL with a different distribution of posts and posters. For comparison purposes, Whittaker et al. (1998) found that on average, 2.9% of newsgroup posters created ¼ of the posts over a 6 month period.

A similar measure is shown in Column 1. of Table 6.5 for each of the DL sets in this study. The values are across the board higher than the 2.9% of Wittaker et al. Only the Large group is close to Whittaker et al.'s value at 3.4%. This may suggest that the health and software support DLs are smaller than the newsgroups examined by Whittaker et al. However, because they logged for twice as long (6 months vs. 3 months) this may also explain the difference.

DL set	1. % of posters creating ¼ of all posts	2. Traffic (posts/day)
All	5.6	16.1
Health	5.1	18.4
Software	8.9	3.1
Large	3.4	25.1

Table 6.5: % of posters creating ¼ of posts and mean traffic for each DL set (3 month period).

Figure 6.7 shows the correlations between lurking levels, posters creating ¼ of all posts and traffic. As the percentage of posters creating ¼ of the posts rises, the level of lurking increases. (R=.448, P<.001). This suggests that lists with high levels of stardom also have decreased lurking levels. This seems counter intuitive, as one might expect that the more dispersed the posting is, the less lurking there would be. DLs with stars may act as magnets to the delurking process.

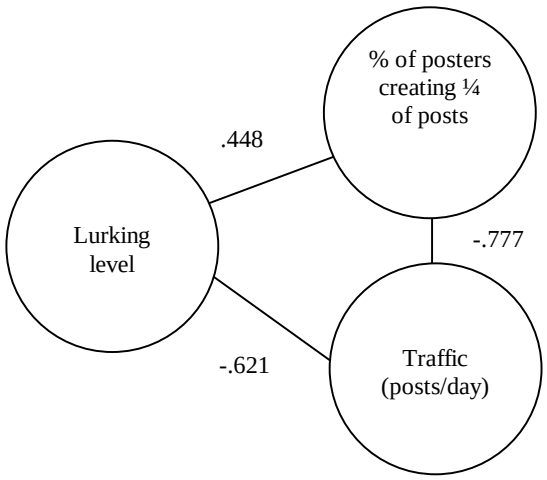

Figure 6.7: Correlation between lurking levels, traffic and % of posters creating ¼ of posts.

Traffic is strongly negatively correlated with both lurking levels and the ¼ of the posts measure (R=-.621, P<.001 and R=-.777, P<.001 respectively). As traffic levels go up lurking goes down and % of posters making ¼ of the post goes down. While there is agreement between these correlations, why there is a strong negative relationship between traffic and the ¼ of the posts measure is not obvious. One might reasonably assume that with increased traffic the posters making up ¼ of the posts would increase, but that is not the case. It is possible that increased traffic has the effect of creating fewer star that also post more frequently. This may be a case of a few shouting the loudest and the most often.

	Question	Snapshot result
R3e	If posting is concentrated with a few posters, how does that affect lurking levels?	The greater the concentration, the less the lurking.

6. Lurking and message length

R3f Are short messages related to lower levels of lurking?

Netiquette (Net etiquette) for newsgroups and DLs suggest that messages be kept short as long messages are less likely to be read (Lehnert, 1998). This is a general recommendation that does not take into account the type of group or the traffic levels. For example, long messages may be appropriate in groups with few posts per day.

Whittaker et al. found that short messages were related to increased interactivity in newsgroups (as measured by thread length). They suggest that long messages may act as a form of "communication overload" (Kraut & Atwell, 1997; Whittaker & Sidner, 1996). In overload situations long messages compete with short ones and are less frequently read because of their length. In a similar manner, DLs with long messages may have higher levels of lurking.

For each message collected the message length was determined (less the header but including any signature). The mean message length was calculated for the DLs and is summarized in Table 6.6. It was found that the mean message length for all DLs was 20 lines. The mean message length data was normalized using the ln function and is reflected in the correlations below.

DL set	Mean message length (lines)
All	19.8
Health	19.0
Software	15.8
Large	33.0

Table 6.6: Mean message length (lines of text less header).

On examining the relationship between mean message length and lurking levels, it was found that mean message length was positively correlated with lurking levels (R=.258, P<.007). As mean message length grew so did the proportion of lurkers. This may suggest that if moderators and posters alike want to draw lurkers into public spaces, then one way to do this is by ensuring that messages are smaller rather than larger. Unfortunately, while the results indicate a correlation, they do not show a cause and effect. Without further work it is unclear whether shorter messages would result in fewer lurkers.

	Question	Snapshot result
R3f	Are short messages related to lower levels of lurking?	Yes: short messages are related to lower levels of lurking.

7. Lurking and clumpiness (threadedness)

R3g If clumpiness is an indication of interaction, does it necessarily follow that increased clump size is related to decreased lurking?

DLs differ from newsgroups and BBSs in their display of complete threads. In newsgroups and BBSs related messages are generally displayed in threads. In email-based DLs, threads are not readily apparent except through common subject headings. Even when a common subject heading is used to organize threads, the messages appear as clumps of messages and show none of the branching associated with news readers or BBS user interfaces. If so inclined, the DL member must manually piece the threads together. For this reason, threading in DLs is better described as clumping of messages.

Whittaker et al. (1998) examined "interactivity" in terms of the depth of conversational threads. They found that the average thread depth was 1.8 messages or almost two additional messages excluding the originating message (i.e., 2.8 messages in a thread).

Table 6.7, shows the mean clump size for each DL set. The health-support DLs have the largest mean clump size (2.5 posts). This is significantly larger than the software-support DLs' mean. The overall thread depth is less than that found by Whitaker et al. (2.4 vs. 2.8). Whittaker et al. also found that 33% of the messages had two or more threads. This being the case, then the DLs examined in the current study are much less threaded than the newsgroups examined by Whitaker et al. One possible explanation is that the lower clump size is the result of email-based DLs having a compromised user interface, which makes following threads more difficult.

DL set	Mean clump size (number of messages)
All	2.4
Health	2.5
Software	2.0
Large	2.0

Table 6.7: Mean clump size (number of messages including originating message).

For the DLs in this study it was found that as interactivity went up (as measured by mean clump size) there was a strong reduction in lurking levels (R=-.631, P<.001 Note: mean clump size data was normalized using ln(x+1)). It was also found that as interactivity went up the mean message length went down (R=-.344. P<.001). Whittaker et al. found a similar relationship. These three relationships are shown in Figure 6.8.

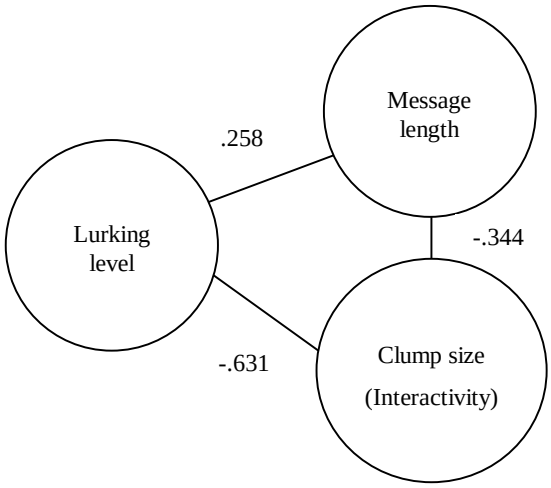

Figure 6.8: The correlations between lurking level, message length and clump size.

	Question	Snapshot result
R3g	If clumpiness is an indication of interaction, does it necessarily follow that increased clump size is related to decreased lurking?	Yes: larger clumps are related to lower levels of lurking.

8. Delurking

R3h Is there a relationship between the number of singleton posters and level of lurking?

Several interviewees in the first study indicated they actively observed how groups responded to members' delurking. Without doing content analysis, it is difficult to determine whether a given post is a form of delurking. The next best measure is a count of how many people post only once. Whittaker et al. describe them as "singleton posters" and found they account for 27% of posts in newsgroups. Singleton posters fall in to two groups: those who receive a public response and those who do not. As in personal relationships, ignoring someone, may be the single most effective way of putting them off further interaction. If a response is not forthcoming, this may be an incentive for lurkers to continue lurking.

The logged messages were examined for singleton posters. Counts were also taken on those who did not receive a response. Table 6.8 summarizes the results. Column 1shows that there were a large number of singleton posters in all DL sets. All DLs in this study have a greater percentage of singleton posters than

that reported by Wittaker et al. One possible explanation for this is that the rate is higher because the current study used a 3 month collection period rather than a 6 month period as in the cited study.

DL set	1. Singleton posters (% of total posts)	2. Singleton posters who do not receive a response (% of total posts)
All	42.0	5.3
Health	36.8	3.6
Software	59.6	11.5
Large	45.1	5.0

Table 6.8: Singleton posters and single posts.

The percentage of singleton posters for health-support DLs is significantly lower than for the software-support set. Knowing how many of these singleton posters receive a response provides a very rough approximation of how well delurking is received. Column 2, of Table 6.8 also shows that the number of posters who did not receive a post is smaller for the health support than the software support DLs (means of 3.6 and 11.5 respectively, and again this is significant). Overall, the proportion of singleton posters who did not receive a response was quite low (mean of 5.3% of posts). This may indicate that singleton posters and perhaps delurkers are noticed and responded to.

There was a strong positive correlation between singleton posters and lurker levels. ($R=.719$, $P<.001$). Similarly, there was a strong positive correlation between lurking levels and the number of singleton posters who do not receive a response ($R=.723$, $P<.001$, see Figure 6.9)

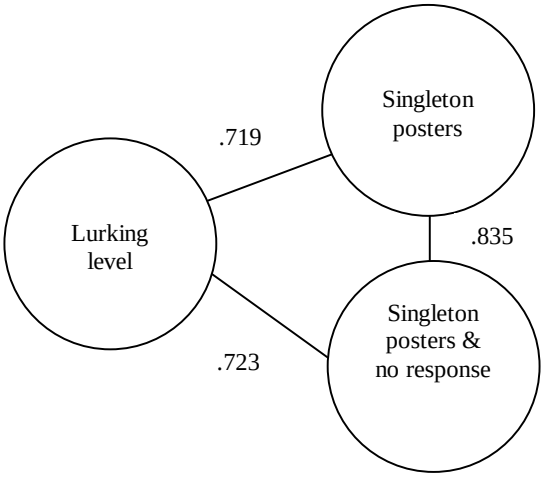

Figure 6.9: Correlations between lurking level, singleton posters (and those receiving no response).

	Question	Snapshot result
R3h	Is there a relationship between the number of singleton posters and level of lurking?	Yes: the more singleton posters (and those who do not receive a response), the greater the lurking.

6.4 Further discussion

Logging posters to understand lurkers and lurking proved to be a fruitful method. Eight areas were investigated and an understanding of some of those areas is starting to take shape (see Table 6.2 for an overview of the results).

Logging has been used by others to test theory, e.g., Wittaker et al. (1998) in their work on mass interaction in newsgroups. In this case, there was very little theory, only questions and a few hunches. For each of the questions we now know whether correlations exist, the strength of the correlations and that the type of group appears to make a difference to lurking. The health and software-support DLs behaved very differently with respect to lurking.

As a follow-up method to the interview study, the logging study proved to be an excellent tool. Part of that excellence comes from the flexibility of applying different kinds of analyses to the logging data. A number of questions asked of the logging data arose from the interview study. For example, it was garnered from

the interviews that lurkers watch the delurking process. This points out how important it is to have the skills available for manipulating logs as this information and its related question were un known prior to the start of the logging study. In the process of doing the analysis, a number of variables were specifically chosen to see if they were related to lurking. All variables were normalized and used as a basis for examining the relationships between lurking level and each of the variables (see Pearson correlation, Table D in Appendix D).

On examining the Pearson correlation table, the number of significant correlations fills more than 80% of the table. More than one-third of the correlations are over .5 in strength. This is a large amount of significant interaction. One of the most surprising findings is that all of the variables were significantly correlated with lurking. Table 6.9 lists the variables ranked according to their R value, i.e., the strength of the correlation. The interactions underlying these correlations need to be better understood before a general model of lurking can be proposed.

Several of the stronger relationships are very interesting. For example clumpiness and traffic have strong negative correlations with lurking level. That means that groups with high levels of interactivity and traffic also have low levels of lurking. This was indeed the case for the health support DLs when compared to the software support DLs. Interviewees in the first study said they looked to see how responsive the group was to delurking. The data suggests that others may do this as well, i.e., lurking level was strong correlated with singleton posters who did not receive a response.

Positive correlations	R*	R*	Negative correlations
No. of lurkers with 0 posts	0.76	-0.66	Posters creating 1/4 of posts (% membership)
Singleton poster – no response (% of posts)	0.72	-0.63	Clump size (messages)
Singleton posters (% of posts)	0.72	-0.62	Traffic- mean posts/day
Poster creating 1/4 of posts (% posters)	0.45	-0.44	Number of topics (clumps)
Number of members	0.33		
Mean message length (lines)	**0.26		

* P<.001, ** P<.007

Table 6.9: Variables with significant correlation to lurking level (0 posts) ranked by R value.

6.5 Summary

The primary goal of this study was to address the question *P3: How many lurkers are there?* If lurking is defined as no posts, then the answer is that lurking occurs much less frequently than the 90% figure reported in the literature. The average level of lurking for all the DLs in this study was 55.5%. In the course of comparing the lurking levels in health and software-support DLs, it was found that health-support groups had significantly lower mean lurking levels (45% vs. 82%). It is not clear from the logging data why this is the case. However, the interview study may suggest a couple of reasons. Persistence may play a roll in keeping the software-support members lurking. They may be a technically more sophisticated group and

understand the negative ramifications better than the health-support members. Another possibility is that members of large lists lurk more frequently. As the software-support DLs were larger in size, this may account for some of the difference. Perhaps the most obvious reason why lurking levels are different is that the types of communities are different and that needs of members in health-support groups are better served through posting. More work is required to clearly understand these issues.

What seemed like a simple question has proven to be a complex web of significant interactions. Further work is required to understand the importance of the interactions and also whether there are cause and effect relationships at play.

As a quantitative follow-up of the interview-based study in Chapter 4 and 5, the logging study proved a capable tool for understanding lurking. There is some irony in studying lurking with a method normally reserved for examining public participation. This work was successful in discovering a number of relationships between lurking levels, DL type, membership levels, traffic and others. Whether they are causal or not, is left to future work.

The data from this study can and may be used for follow-up work, e.g., to determine whether lurking is related in any way to the diversity of topics within a DL (i.e., breadth vs. depth of the DL). Another area worth pursuing, but perhaps outside of this data set, is the investigation of high-traffic DLs and their members. For example, how do members cope with high traffic levels?

In the next chapter, a discussion of the results from both studies will be presented. Accompanying this discussion will be a review of the methods used in the two studies and also a new definition for the term, lurk.

Chapter 7: Discussion

Overview

- Findings from the two studies are discussed.
- Design implications related to DLs are outlined.
- Methods used in the two studies are reviewed.
- An improved definition of lurker is proposed.
- Discussion is summarized.

Much of the discussion of the results has already taken place in Chapters 5 and 6.This chapter opens with a discussion of five results which are of special interest. This is followed by a series of design implications that should be of interest to software and community developers. The interview and logging methods employed are then reviewed with an eye to understanding how they contributed to the results. The second to last section proposes an improved definition of the term, *lurker*. The chapter ends with a summary of lurkers and their lurking.

7.1 Highlights on lurkers

Several results stand out from the two studies and will be discussed in this section:

- lurkers are not free-riders
- three models of lurking
- lower lurking levels in health-support DLs
- sense of community
- high traffic correlates with low lurking

Lurkers are not free-riders

The subject of lurking has strong universal appeal to researchers and non-researchers alike. That appeal no doubt comes from the fact that each and every one of us has either lurked, is lurking or will lurk in the future. As one researcher said, “you’ve got it wrong, lurking is normal, it is the people who post who are abnormal” (B. Wellman, personal communication, May 1999). Taking that perspective it is difficult to equate lurkers with free-riders.

Chapter 2 mentioned that Kollock and Smith (1996) describe lurkers as free-riders. Describing lurkers as free-riders classifies them for their lack of public participation and their use of resources without giving back to the group. Even when lurking is narrowly defined, e.g., one or fewer post/month, the vast majority of DL members in the second study are lurkers (81%). This being the case, how do online groups survive in the face of almost universal free-riding?

One explanation is that lurking is not free-riding, but a form of participation that is both acceptable and beneficial to online groups. Public posting is but one way in which an online group can benefit from its members. Members of a group are part of a large social milieu, and value derived from belonging to a group may have far-reaching consequences. For example, information supplied in health-support groups may end up enlightening a member. When the group member then uses that information to seek better medical care, physicians and other health-care professionals also benefit from this knowledge. The online group is but one way in which the member communicates with others. Online groups are but one of many places for interaction, and although it may not seem like it from a research perspective, life for most members is more than life in the online group.

A second explanation is that a resource-constrained perspective may not apply to online groups where the centralized cost of servicing 100 members isn't much different from that of serving 1000, or even 10,000. In large DLs the danger could be in not having enough lurkers. If everyone posted in large DLs, there would be a flood of messages that could make interaction very difficult.

Based on the results from the two studies, the following suggest why lurkers are not free-riders:

- **Lurkers work at knowing the group:** Interviewees described putting substantial effort into understanding a group. This work was a benefit to both the group and the lurker. It provided the lurker with the knowledge of whether the group was a good fit. It also provided them with an understanding of the social dynamics of the group. These are important consideration taken by the lurker.
- **Lurkers try not to add to the chaos:** Many of the interviewees found that groups can be chaotic environments. Examples of this chaos include high levels of posts, duplicate posts by different authors, and irrelevant or inflammatory comments. Lurking was a form of participation that does not add to the chaos.
- **Lurkers are often well connected:** Interviewees generally described groups in which the topics were of interest to them. Often these groups were related to other aspects of their life, e.g., one of the interviewees had a passion for history and as a result joined a number of military history groups. Some interviewees joined groups based on friends recommending they join. Belonging to an online group is just one expression of a web of related activities. Contributions can and do take place outside of a group's public space.
- **Some lurkers side post:** Several interviewees said they made connections to individuals outside of the online groups. These connections are a valid form of communication and have value for the individuals and thus the group as whole. Reasons for side posting varied and include a desire to contact individuals rather than the whole group, not wanting to get involved in a public dialogue when time was of the essence, and feeling more comfortable in one-on-one emails.
- **Lurkers make a commitment:** Joining a DL appears to be somehow different from browsing a BBS or a newsgroup. For several interviewees, the process of joining and either explicitly or implicitly agreeing to the rules of the group is a form of commitment. For these interviewees, making a commitment to join a DL and then lurking did not feel in any way like free-riding. Rather, it was a way of assessing a group and determining the group's value to the member and the member's value to the group.

Models of lurking

The three models of lurking (*filter*, *gratification* and *persistence*) formulated in Chapter 5 encapsulate lurkers' processes, needs and circumstances. The *filter* model describes the relationship between members in terms of a series of filters or barriers. If any one of these filters comes into action, then posting does not occur and the member lurks. The *gratification* model describes the meeting of one's needs in the most efficient way possible. Lurking, at least in some situations, is the least costly way of getting one's needs met. The *persistence* model is less about interaction and more about how the properties of the environment have an effect on lurking. In the case of lurking, persistence in the form of persistent messages has three primary effects. Persistent messages helps the lurker lurk, gives cause to lurk, and makes work for the lurker. It is clear that further work is needed to provide a unified model of lurkers and that these elements will likely play a part in building that model.

Lower lurking levels in health-support DLs:

Lurking levels in health-support DLs were found to be on average significantly lower than in software-support DLs. Although this work does not present evidence for why this is the case, there are several possibilities. Members of technical support communities may be more aware of the issues related to persistence, e.g., messages on the Net have an indefinite life and are searchable. This being the case,

technically sophisticated group members may be more wary of posting. Technical support members may also have affiliations with institutions or commercial enterprises and may need to protect their own identity or that of their company. Lurking may also be a way of not showing their ignorance. While it is not clear this is the case, gender may play a role as software-support groups may be male dominated. Other researchers have shown that highly interactive groups tend to have more women in them (e.g., Roberts (1998))

The needs of the members of health and software-support groups are likely very different. As shown in other studies (Preece, 1998; Preece & Ghozati, 1998), health support communities have high levels of empathy. Interaction, in the form of receiving and giving empathy is probably higher in the health-support DLs. As was shown in Chapter 6, high levels of interactivity (threading/clumping of messages) were related to lower lurking levels. Relationships may be more personal in the health-support groups and there may be less flaming and hostility. Both could act to reduce or eliminate some of the barriers to public participation.

Sense of community

Several interviewees said they felt a sense of community while lurking. This is curious as it flies in the face of what many consider to be the defining elements of community. Definitions of community commonly incorporate the following (e.g., Erickson, 1997; Preece, 2000; Roberts, 1998; Wellman & Gulia, 1999; Whittaker et al., 1997):

- notion of membership
- relationships between members
- commitment and reciprocity
- shared values
- collective goods
- duration

To understand how lurkers can have a sense of community, each of the above attributes is examined from the perspective of the lurker. The underlying assumption in this discussion is that online groups may be communities.

Notion of membership: Interviewees were members of the groups they discussed. This was demonstrated by their knowledge of the community and the effort they put toward learning about the community. Belonging to a community is often a process of coming to know the members, traditions, rules and language. Interviewees mentioned this process and also mentioned that lurking was a way in which they learned about the group and eventually considered themselves to be members. It may be that for some people and at certain stages of membership (e.g., being a new member) lurking is an indirect way of saying they are not yet members, but are trying to be.

Relationship between members: Forming personal relationships with community members was important to some of the interviewees in some of their communities. For other interviewees, becoming a member of the group and forming a personal relationship wasn't necessarily part of the interviewees' desires or needs. For example, finding out a piece of information did not have to incorporate the development of a personal relationship. It could, but wasn't required. In groups where the interviewee's goal was strictly information, then a sense of community was not felt. However, where personal relationships were pursued, often outside of the public space (e.g., through email), then a sense of community was possible. It is also possible that a lurker can feel they know someone very well from their public postings and in that way feel kinship with that person. One interviewee mentioned that stories related to members' health issues were particularly powerful in engendering a sense of community. The non-reciprocal relationship of the story teller and the lurker provides a sense of community for the lurker, even if it bends the concept of communities being reciprocal in nature.

Commitment and reciprocity: It was obvious that the interviewees with the sense of community were very committed to the community(s). This was shown in their effort to understand the community, often through the careful reading of messages and side posting to members. Many lurkers are willing to support individuals in their dialogue outside of the public spaces. For example, Katz (1998) experienced a deluge of private and supportive responses from lurkers when he was verbally lambasted in a BBS.

Shared values: Most of the communities mentioned by interviewees are topic based. These topics draw interested parties into them, either by sharing or becoming familiar with a common set of values, knowledge, or practices. This effort expended in becoming knowledgeable about a group is in a sense a measure of the respect for the community. In health-support groups, members' dialogue and stories allow other members to share in their experience and identify with the authors' (Preece, 1998).

Collective goods: It is unclear how lurkers contribute to the collective goods of the community. It may be that their contribution lies outside of the public dialogues, e.g., in other ways such as sharing their experience with others outside of the online community. They may spread the word and act to enlarge the community by drawing in new members. This broad interpretation of goods includes the community itself and the persistent dialogue, i.e., resources and information.

In noisy or chaotic groups, lurking allows the collective goods, i.e., the already existing messages and dialogue, to be more easily perceived by the whole group. Interviewees were aware that public participation was not always good for the group. Several interviewees said they knew others would voice similar views and adding a message to the dialogue would not add value to the discussion.

Duration: As already mentioned, the interviewees were committed to understanding the community and spent a considerable period becoming familiar with the community and following the conversations.

While it is a bit of a stretch to say that the lurkers met all the criteria for being members of the community, some nonetheless had a sense of community. For them, having a sense of community was likely different from them being members of the community. Even they would probably make that differentiation.

High traffic correlates with low lurking

During the first study (Chapter 4), interviewees described the effort required to manage DL traffic. If there were few messages, then the DL was effectively out of mind and required little or no effort. If there were many messages, then the DL became burdensome. Several interviewees cited newsgroups as being less useful because of the large volume of messages. They also mentioned that message quality was very important, e.g., content, knowledge base of participants, and courtesy. Several interviewees left newsgroups because of high traffic rates and poor quality of messages. In contrast, results from the logging study tell a different story. As traffic levels rise, lurking levels decrease ($R=-.631$, $P<.001$).

The mean traffic level of the health-support DLs was 18.4 messages/day, and one DL exceeded an average of 97 posts/day. These higher-than-expected numbers suggest that these DLs are somehow different than the DLs participants described as being ideal in the interview study, i.e., less than 6 messages/day. Why the discrepancy? It is possible that these DLs supply such high-quality content that their members are willing to put in the higher effort to deal with the large volume. It is also possible that high traffic DLs act like many little DLs, each identifiable by a set of subjects and/or authors. The observed high traffic levels suggest that what is an acceptable and perhaps necessary traffic level in one DL may be unacceptable in another. It also suggests that motivation, in addition to quality of messages, is an important facet of acceptable traffic levels.

Understanding what constitutes acceptable traffic rates is an important issue in designing online communities. E-commerce is already running into this problem. For example, when sending promotional materials through distribution lists, it is important to understand how much email can be sent before customers perceive it as a nuisance. Understanding how DL members cope with and make use of high volumes of messages is important for the designers of email-client software. Lastly, messages from DLs are

not received in a vacuum; they compete with messages from a variety of other sources, including personal and professional correspondence, and email from other DLs.

7.2 Design implications

As a means for asynchronous group communication, DLs have gained wide acceptance. This is in part due to their use of the most ubiquitous of Internet tools, the email client. A number of issues and design implications are illustrated in this section. These implications have been drawn from both the interview and the logging studies. One area that is not covered here is the work currently being done on visualization of presence (e.g., Donath et al., 1999; Viegas & Donath, 1999). Presence and its effect on lurking will be examined in the next chapter on future work. The five areas below are included in the following discussion.

- email client
- server software and administration
- supporting information
- alternative access mechanisms
- member

Email client: There are two leading ways in which all email clients can be improved: by showing threading, and improving filtering. Threading provides lurkers with the ability to judge whether messages are valuable, and how to deal with them. It also allows the user to follow conversations more easily. At this time, threading is poorly implemented on most email clients. Filtering has the capability of separating and thus organizing multiple DLs into separate areas and thus reducing clutter in the inbox. Filtering is readily available on most GUI based email clients, but is not frequently used.

Server software and administration: At the server level, several improvements could make life easier for the lurker. However, some improvements negatively affect other areas. For instance, digests are intended to reduce inbox clutter, especially in high volume DLs. Unfortunately, thread following is compromised when messages are delivered in digest form. DLs that send digests to new subscribers may also have higher levels of lurking (Nonnecke & Preece, 2000a). Whether this is a result of digests being less well read, more difficult to respond to, or making threads harder to follow is unknown.

Many DLs add a prefix to the subject header as a means of identifying messages. These prefixes may make DL identification easier, but likely obscure the actual subject header. It is unclear whether knowing that a message comes from a particular DL is more important than the subject. It will likely depend on many factors, including the volume of messages in the inbox, the rate of receipt, and purpose of belonging to the DL.

Most DL administrators prevent access to membership lists. More often than not they also prevent messages being broadcast by non-members. There is however, very little they can do to prevent the pilfering of addresses from archives. Some members have taken up the anti-spam challenge by supplying incorrect return addresses.

Supporting information: At the level of supporting the lurker with information related to the DL, providing an accessible, current, and usable set of information is important. Creating links to it in all outgoing messages would provide access. Within the Web site or wherever it may reside, access to an archive is an important information resource for many lurkers, particularly if they are trying to understand the nature of the DL, or looking for specific information. A usable interface should allow lurkers to browse, follow threads and search for information.

Alternate access mechanism: Alternatives to email clients are Web-based UIs to DLs. Web-based UIs can either reduce or eliminate reliance on the email client. In doing so, many of the problems described so far could be reduced, e.g., threading is usually apparent in Web-based interfaces, anonymity is frequently built into the system, thus ensuring safety and privacy. In addition, an archive, search tools and supporting information can be integrated into the environment.

On the down side, the user may have a different identity and potentially a different password for each DL. Also, UIs may differ between DLs. This would require familiarization with each different UI. In contrast, email-based DLs utilize a single familiar UI of the email client. In addition, the email client can receive email automatically, whereas, Web-based DLs rely on the user to seek them out.

For the security conscious, the use of Web-based DLs may offer a preferred solution. However, it is unclear whether security is more important than the convenient and consistent albeit somewhat underused and noisy UI of the email client.

Member: Email clients are facing functionality bloat already. As a result, adding additional functionality for lurking may not be the best approach. Improved lurking may come through improving the skills of the lurker. While the current email clients may not have been specifically designed with lurking and DLs in mind, many of their facilities go unused by the lurker, e.g., use of secondary mail boxes and filtering. This is in part due to users not being familiar with the functionality of the email client, but also stems from the way in which they view and use the inbox as a central repository. As is the case with other software, DL members use the tool to the extent that fulfills their immediate needs. It's likely that the level of participation (posting or not) in one or more DLs is a function of their skills in using the email client. Other factors will also be at work, such as volume of email, personal strategies, motivation, time available, etc. By improving the understanding of strategies and the context in which the strategies develop, tools can be improved and a better model of the lurking developed.

It is not surprising that some of the lurkers' goals lead to contradictory design implications. On one hand, privacy is a very important issue, and on the other, participants desired more information about the DL and its members. Email clients are relatively simple, well understood tools which in part accounts for the popularity of DLs. Improving their facility as DL front ends may increase their complexity and thus compromise their broad appeal. Full featured Web based UIs compared to DLs hold promise for eliminating many of the problems associated with the email-based UI. Whether DL members are willing to switch to an alternative UI is unknown.

7.3 Evaluation of the methods

In recent studies of online groups, the primary source of information has come principally from participants who actively conversed in the online groups, i.e., from those who were readily observable. These studies have covered broad areas, including the nature of online communities (Wellman, 1997), the development of friendship (Park & Floyd, 1996), the role of empathy in group discussions (Preece, 1998), and the differences between men and women (Roberts, 1998). Work has also been done on specific kinds of online communities, e.g., therapy (King, 1994), education (Hiltz, 1993), business (Sproull, 1986), and health support (Preece & Ghozati, 1998). While knowledge is growing, it is nevertheless a selective knowledge based on observations of those who post. Unlike the studies mentioned above, this thesis has focused on those who do not participate publicly, the lurkers. Two methods were used to study lurking in this thesis: semi-structured interviews and logging. The remainder of this section will discuss these methods and their value in understanding lurkers.

Semi-structured interview

The first method employed in studying lurkers was the semi-structured interview. This was a qualitative study with the primary goals of understanding why lurkers lurk and what lurkers do. This method sought to elicit information about lurkers from group members. The assumption being that interviewees would be lurkers.

By using a semi-structured interview, the interviewer has the ability to follow and explore new ideas and information as it is garnered from the interviewee. An initial structure was used and improved upon in succeeding interviews. A consequence of this approach (and many other approaches) is that the initial position is elastic and can change during the course of a specific interview and between interviews. Initial

assumption can be quickly verified or disproved and new paths of understanding open up. Each successive interview incorporates the knowledge from the interviews before.

Several things became very obvious after the first couple of interviews. For one, the semi-structured interviews worked very well. By emphasizing the online groups and then having the interviewee describe their participation in the group, it was possible to understand lurking, not just in terms of lurking behaviour, but also in terms of participation in general. Understanding the types of groups they joined, the reasons they joined, their activities and duration of their membership painted a very rich picture of their lurking and participation. For many interviewees, talking about a particular group experience proved to be a good anchor for their discussion.

A number of interviewees had a very good understanding of why and how they participated. For these interviewees, it was often easier for them to describe their strategies and then give examples using a specific group. These interviewees appeared to be more reflective and less reactive. Their approach to explaining their knowledge and experience while different from the anticipated interview structure, in no way lessened the value of their input. This leads to a discussion of a bias that was unintentionally introduced into this work.

One conclusion drawn from the study is that lurking is a strategic and idiosyncratic activity. This conclusion may be a result of the interviewees being well educated and comfortable in talking about their use of the technology and how it affects them. The interviewees were in all likelihood more comfortable with the technology than the average online group member. If a less technically literate set of interviewees had been used, then the results would likely be biased in some other way. For example, studying lurking habits of teenagers would likely lead to some very different results. They may be much more adventurous in their use of technology and would probably seek different kinds of interaction.

The collection of specific quantifiable information was less important than the exploration of the issues during the interview. Given that this first study was not intended to carry the burden of quantifying lurkers and their behaviours, this was a reasonable approach. A simple count did show that everyone was a lurker. Some interviewees lurked all the time, some lurked in specific DLs and others lurked at specific times. Knowing that lurking was prevalent was an important finding.

The second quantitative and important piece of information gained from the interviews has two parts. The first part is that DLs were the predominant online group and second, that lurking levels within DLs were high. These interview findings reinforced the need to further understand lurking and showed that DLs would be a good place to start looking. This set the stage for the log-based study of DLs.

The semi-structured interview provided a large amount of information. Making sense of that information proved to be one of the most valuable parts of the process. The three models of lurking (*filter*, *gratification*, and *persistence* models) put forth in Chapter 5 reflect the richness of the interviewee information and this researcher's need to make sense of it. This was an iterative process, one in which rationales and activities were aggregated along many different dimensions.

The interviewees' stories of their group participation hold potential for describing lurking in a way that could make the lurking experience more accessible to community developers and tool designers. Given the dearth of lurker information for grounding group and community design, these stories could provide a means of establishing and fostering environments suited for lurking.

While the interviews proved informative in answering two *primary* questions, they also provided input into the logging study. A group of questions arising from the first study were used to direct the analyses in the second study. This synergy between studies is not one way. For example, now that correlations are shown to exist, interviews can be used to understand whether there are specific cause and effect relationships. In addition, both of these studies can contribute to other types of studies, e.g., Web-based questionnaires for investigating posting patterns outside of the public spaces.

Logging DL messages

For all questions (see Table 7.1), the logging data provided a new understanding of lurking. If there is a significant issue with the results it is that they do not address whether the relationships discovered are cause and effect. That will require further investigation with perhaps quite different methods.

	Questions asked of the logging data	Finding
P3	How many lurkers are there?	Fewer than expected: still high with an average of over 55% for all DLs (when defined as 0 posts in 3 months).
R3a	Does lurking in health and software-support DLs differ?	Yes: health-support groups have lower levels of lurking (45% vs. 82%).
R3b	If lurking is defined as no posting, what happens to lurking levels when the definition is broadened to include minimal levels of posting, e.g., 1 post/month?	Lurking increases rapidly and then levels off as definition is broadened. Health-support groups maintain their lower levels of lurking (75% vs. 97% for software when lurking is defined as 3 or fewer posts/3 months).
R3c	Is there a relationship between lurking and the number of members in the DL?	Yes: smaller DLs have fewer lurkers.
R3d	Is there a relationship between lurking and the traffic level of the DL?	Yes: higher traffic means lower lurking.
R3e	If posting is concentrated with a few posters, how does that affect lurking levels?	The greater the concentration, the less the lurking.
R3f	Are short messages related to lower levels of lurking?	Yes: short messages are related to lower levels of lurking.
R3g	If clumpiness is an indication of interaction, does it necessarily follow that increased clump size is related to decreased lurking?	Yes: larger clumps are related to lower levels of lurking.
R3h	Is there a relationship between the number of singleton posters (and singleton posters who do not receive a response) and level of lurking?	Yes: as the number of singleton posters rises (and those who do not receive a response), so does the lurking.

Table 7.1: Questions put to the log-based study and the findings.

The demographic study collected messages from 109 DLs. Message logging can be employed on a much larger scale, e.g., thousands of groups can be logged. Having studies which incorporate many more DLs would allow other questions to be investigated. For example, are there differences in lurking levels between groups with the same topic but of different size.

One of the primary advantages of logging is its low data collection costs. Doubling the number of DLs in this study would have involved very little extra cost. The only costs would have been in joining and maintaining the membership. For example, some DLs renew memberships every three months or so. Failure to reply to the renewal notifications results in a loss of membership. Another advantage logging has is that it can be carried out over a variety of time periods. The current study was 3 months in duration, however, it could just as easily have been carried out over 6 months or a year with very little additional cost.

If there is a flaw in using logging to study lurkers it lies in the fact that the DL membership fluctuates. Of the 109 groups surveyed, only 3 allowed access to the group member list. The membership in these three lists were followed for several weeks and showed very little change. Unfortunately this may not extrapolate to all the DLs in the survey. Even if an accurate accounting of membership were possible, knowing how to

interpret that information is problematic. For example, should a member be counted who enters a DL for a week and then leaves. Even when there is some assurance of consistent membership, there is still the problem of knowing whether that member is there in name only. For example, they may have joined and then changed jobs and thus their email address. They have then joined the group using the new email address. They could be counted as a lurker under one email address and a poster under the other. It is under these uncertainties that the logging study and its results and conclusions must be understood. This work is very much an approximation of a dynamic and somewhat uncertain system.

7.4 Improved definition

A substantial gain in knowledge has occurred as a result of the work contained in this thesis. As mentioned in Chapter 1, the online Jargon Dictionary (Jargon Dictionary, 1999) provides the following definition for the term, *lurker*:

> One of the 'silent majority' in a electronic forum; one who posts occasionally or not at all but is known to read the group's postings regularly. This term is not pejorative and indeed is casually used reflexively: "Oh, I'm just lurking." When a lurker speaks up for the first time, this is called 'delurking'.

The definition asserts the following:

- lurkers are silent
- lurkers are in the majority
- lurkers post occasionally or not at all
- lurkers are known to read postings regularly
- the term is not pejorative

In order to improve upon the definition, each one of the assertions will be examined. New information will be added where appropriate and a new definition will be put forward.

Lurkers are silent: Lurkers silence, if it can be called that, is an idiosyncratic and strategic set of processes that involves more than being publicly silent. As discovered in the interviews, lurkers do participate through private exchanges outside of their online groups. Their silence is limited to their lack of public posting. In asynchronous environments like DLs, BBSs and newsgroups, the silence is unnoticeable by the group. In these environments, lurkers do not have a presence in the way they do in a MOO. In large groups, lurkers' silence may help reduce the chaos and thus keep the group more manageable. This is particularly true of DLs where messages arrive in a general purpose email client and compete for attention with non-DL messages. In BBSs and newsgroups this is less of an issue as these employ single purpose tools, i.e., BBS Web UI or a newsreader.

Lurkers are in the majority: It depends. As was seen for at least some of the health-support DLs, lurkers are in the minority. Lurking levels vary widely, not only between groups but sometimes within a group. In groups with spikes of activity, whole memberships can change from lurking to posting and back again in a very short period of time. The use of the term "silent majority" has a strong political connotation, one in which the majority is somehow unified. This is clearly not the case when it comes to lurkers.

Lurkers post occasionally or not at all: This is the heart of this definition. It looks at participation from a numbers perspective, one in which lurking is defined in terms of the volume of public posting. Defining lurking as a function of a posting rate skews the definition. There is no measure of the quality of the post or its ability to engender discussion, and implies that all public posts are equally valuable to a group. This is not the case. For example, a member may be able to contribute questions to a group, but be unable to publicly respond to a query. This person could be called a response lurker. Yet under the current definition of lurker, this person is not a lurker. This part of the definition doesn't account for the value that a poster might have in developing or sustaining interactivity. For example, a strong public post by a lurker could create an ongoing dialogue well after the initial post. As was seen in the log-based study of Chapter 6,

increased interactivity was strongly correlated with lowered lurking levels. Knowing the factors involved in creating and sustaining interactivity in a group could provide insight into lurkers and lurking.

Lurkers are known to read postings regularly: This may or may not be the case. The *filter* model discussed in Chapter 5 describes 4 filters which can act as barriers to public posting. These same filters can act to limit reading and other forms of participation. To describe lurking in terms of reading and not the other activities, e.g., side posting, or any form of off-line participation, is limiting. It is limiting not only in terms of understanding lurkers, but also in the value that lurkers can have to the group and that the group can have for lurkers. For example, knowing that lurkers will often expend considerable effort in trying to understand a group, suggests that sources outside of messages should be made available to new members. A good example of this in health-support communities, where at least one interviewee indicated they sought out members' stories of their own health problems in order to understand both the health problem and the community's perspective on the issues.

The term is not pejorative: The origin of the word lurk suggests it is pejorative. For the most part interviewees thought of the term lurker as pejorative. In many online groups, it is a frequently stated sentiment that lurkers are not pulling their weight, i.e., they are thought of as what Kollock and Smith (1996) call "free-riders". Based on the interviews, lurking is a strategic, active and beneficial form of participation. For example, lurking can occur when members are new to a group. Having them participate publicly without understanding or being comfortable with the group would not be beneficial to most groups.

Calls by moderators and others for increased participation by lurkers may be futile without the moderators themselves reflecting on why lurkers do not participate. It is likely that the quality of moderation, the topic of the group, and the tone of the dialogue have more than a little to do with the silence. A number of authors have pointed out the critical role of moderation in online groups (Berge, 1992; Collins & Berge, 1997). Rather than blame lurkers for their silence, group participants, moderators, owners and managers need to assess the role of the group and how the group achieves that role. For example, is the group large enough to develop and sustain dialogue, and do the current members encourage dialogue in the way they publicly communicate with one another.

New definition: Creating a new definition is an interesting task, given the complexity of lurking . The following blends many of the above comments of the previous definition and is a first step in providing a realistic and accurate definition of lurking.

Condensed definition of lurker: Anyone in an online asynchronous forum who for any number of reasons chooses not to participate in public. Lurker participation is idiosyncratic, situated, and may be strategic and very active. The term is frequently used pejoratively.

Extended definition of lurker: The term, *lurker*, is frequently used pejoaratively and usually refers to anyone who never posts or posts infrequently. In fact, lurking is non-public participation. Lurking is a situated action, and many personal and group-, work-, and tool-related factors affect the activities and level of public and non-public participation. Lurking is "normal" in the sense that everyone is likely to be a lurker at some point in time. Lurkers are heterogeneous in most respects except in their lack of public posting. Therefore, in the absence of an understanding of the context in which it takes place, *lurker* is a meaningless term. Avoidance of the term *lurker* is recommended. Instead, the term *non-public participant (NPP)* is suggested. *NPP* is not pejorative and suggests there are other forms of valid participation outside of public posting.

7.5 Summary

In the chapters leading up to this chapter, two studies were described, results presented and the results from each of the studies were discussed. This chapter discussed different areas that integrate the results from the two studies:

- highlights from the two studies
- design implications for DLs
- evaluation of the methods used in the two studies
- improved definition is proposed

Five different highlights were discussed in this chapter. The first was whether lurkers can be called free-riders based on the evidence. The answer is no. They do not meet the free-rider criteria. Models of lurking were reviewed in the second highlight. These three models provide insight into how lurking can be perceived as more than reading messages and not posting. Each model shows a different aspect of lurking, i.e., barriers to public participation, getting one's needs met efficiently, and the effect of persistent messages. The third examined why health-support DLs have lower levels of lurking. There are several possible reasons that include the very nature of health-support groups and the way in which members communicate in them. Another reason may be that members of health-support groups are less aware of the issues surrounding persistent messages and as a result may be less inhibited in their public posting. The fourth highlight discussed how lurkers can have a sense of community. This discussion was based on whether lurkers met the criteria of being community members. In the end, it is not so much that they are community members, but that they have a "sense" of community. The fifth highlight is the interesting result from the logging study, i.e., that as traffic levels go up, lurking levels go down. This result flies in the face of the feedback from the interviews which suggested that lurking is more likely to occur in high traffic lists. Obviously, something else is at work. Several possibilities are suggested.

In the second section, a series of design implications were put forward. These implications cover email clients, DL server software administration, supporting information for DL members, access mechanisms other than email clients, and changes at the member level. These design implications will be of value to DL members, moderators, designers, administrators and builders of online groups.

The methods used to study lurking were reviewed in the third section. The method of using a small group of participants and interviewing them with regard to the membership and practices within online groups was an effective technique for exposing a wide variety of issues related to lurking. This in turn fed into the logging study which provided a wealth of quantitative information about lurking in DLs. As a result of this work, knowledge of lurking has been improved and expanded upon.

The last part of the discussion deals with an improved definition of the term, lurk. The new definition incorporates the knowledge of lurking gained in this thesis. Improving the definition will be an ongoing process as more becomes known about lurking.

The next chapter contains a critique of the methodology, a discussion of future work, and conclusions.

Chapter 8: Critique, future work & conclusions

Overview

- Critique of the methodology is presented
- Future work is discussed
- Conclusions are put forward.

This chapter begins with a critique of the methodology. Future work is then discussed and the chapter and thesis close with a series of conclusions.

8.1 Critique

This thesis employed two very different methods to address a range of questions on lurking. The two research methods were chosen to provide qualitative and quantitative information and supply a degree of triangulation. Triangulation of quantitative and qualitative methods is important in establishing complimentary perspectives and evidence (Sudweeks & Simoff, 1999).

The interviews were to provide qualitative information on why lurkers lurk and what lurkers do. The interview data resulted in 3 models of lurking, 5 lurking strategies, and 33 groups of reasons for lurking. The study also provided quantitative information indicating that lurking was common amongst the interviewees, that DLs were the most popular online forum, and that lurking was common in DLs. This study also raised a number of questions that were useful in probing the data from the second study, most notably, a question related to acceptable traffic rates.

While extremely successful at uncovering useful information on lurkers and lurking, there are a couple of caveats. The first is that the sample size was relatively small. This worked both for and against the results. With a small number of participants, data analysis is manageable. Larger numbers of participants may have resulted in other quantitative measures, e.g., counting types of groups. There are also tradeoffs in terms of how the actual interviews could be managed. A large number of interviews would require more structure and have less flexibility. Otherwise, the volume and complexity of data would be difficult to deal with. The end result would be a decrease in both the breadth and depth of the interviews. This brings up the second caveat. The results are very likely biased as the interviewees were well educated and relatively sophisticated users of the Internet. How the bias shows up in the results is unclear. Further work is required.

The primary role of the logging study was to provide information on how many lurker there are. In doing so, 109 DLs were logged over a 3 month period. The 150,000 messages logged became the basis for the demographic analysis. All this occurred at very little cost. Logs of this type can provide more than just quantitative information, e.g., they can be used for discourse analysis. In this study, they were used to characterize the data along fairly simple and quantifiable dimensions, e.g., traffic rates, DL size, and lurker rates.

The logs proved to be an excellent information source and were used to address 9 specific questions. Out of this work came an understanding of differences in lurking between health and software-support DLs and also of correlations between lurking levels and other variables, e.g., mean message length. The log study did not, nor was it expected to show cause and effect relationships. That is left to future studies.

Carrying out this type of study requires relatively sophisticated programming skills. If they had not been readily available, this work would not have been possible. It should be noted that the analysis of the logs

assumed constant membership in the DLs. This is clearly not the case as DL membership does change over time. To understand how changing membership might affect the results, a number of DLs were closely observed over a period of weeks. Membership changes were small and would not lead to different end results. It is under these uncertainties that the logging study and its results and conclusions must be understood. This work is an approximate description of a dynamic and somewhat uncertain system.

8.2 Future work

Given that this is a relatively untapped research area, this thesis has unleashed a wealth of possibilities for future work. The following areas will be discussed in this final section:

- current logging data
- other logging studies
- effect of showing presence
- visibility of individual
- design implications
- individuals and individual groups
- other group environments

Current logging data

Without going to further studies, the current data has more to offer. For example, lurking may not be a continuous state and could be punctuated by periods of public posting based on topic or need. Using the current data set, there is no reason why analyses of this type cannot be carried out in the future. The raw data could also be used from an ethnographic perspective, one in which content and dialogue analyses could be carried out. Examples of these kinds of analyses can be found in Preece and Ghozati (1998), and Worth and Patrick (1997).

Other logging studies

One area where logging has a future is in working with larger sets of data that are collected and at least partially analyzed in a more automated way. Many of the techniques used here could be used to monitor thousands of groups as seen in other studies (Smith, 2000). Groups could be selected not just on their topic type, but on size, interactivity or some other aspect. It would then be possible to understand the dynamics of the groups from multiple perspectives. For example, it might be that the single most important driving momentum behind a group is not the topic or its size, but one or two of its most vocal members. There are a number of areas where further logging studies will be of value:

- high traffic DLs and how the group and its members deal with it
- how different DL topic types relate to lurking
- stardom and its relationship to group health
- group/community health and lurking, interactivity, membership turnover, etc.
- size and its relationship to lurking
- examination of initiating posts vs. responses
- gender difference especially in high traffic groups

Effect of showing presence

Work being done on showing presence in groups (Ackerman & Starr, 1996; Donath et al., 1999; Viegas & Donath, 1999) may very well change lurking and the anonymity of lurking as it exists in DLs, BBSs and newsgroups. If new group environments start showing the history of individuals or naming individuals in the process, then lurking will become very different. If lurking is still a desired part of participation, and anonymity is an important part of that process, then work arounds will need to be developed to accommodate lurking. For example, lurkers could participate using aliases or perhaps some sort of optional

cloaking could be incorporated into these environments. It is important for the researchers and designers of these systems to consider the impact of making members more visible, especially at the individual level. Areas where changes in group environments would have less of a direct effect on lurking include information at the group level. For example, a system that shows what percentage of the group read a post could be useful for new members in identifying important posts and posters, and thus come up to speed on the group more quickly. Change in UI to show presence may very well change the landscape for lurking. Knowing how this may affect lurking is an interesting problem.

Visibility of individuals

A posting rate of one post/month is an infrequent level of posting. It could be argued that most of what is being done by members at this level is not posting. Presence or visibility of members within a list may be a better indicator of lurking, i.e., is a member known to the other group members in a way that makes them somehow recognizable and thus not lurkers. Defining lurking as a function of the visibility of the poster suggests that other factors would influence this visibility, e.g., the number of members, the number of posters, the activity of the list, and the value and/or notoriety of each participant. It is possible that someone who flames on an irregular basis may be seen as less of a lurker than someone who contributes in a regular but less visible manner. Understanding the visibility of individuals within a group is a topic related to lurking and would be a good follow-up to the work in this thesis.

Design implications

The current studies suggest design implications for DLs. Each of the design implications could form the basis for further research. For example, threading is poorly shown in email-clients. Adding threading functionality to email clients could provide an enriched environment which could change the way a person deals with their DL email. Another example concerns the use of filters. While they exist in most email clients, they are not extensively used. Would educating users on their use have an effect on their lurking?

Individuals and individual groups

Understanding lurking will be incomplete without further studies of individuals and individual groups. Longitudinal studies of an ethnographic type could provide a wealth of information that would be strongly tied to the context of the lurking and participation in general. Work has been done in this area, but has focused on participation in the online groups (e.g., Mason, 1999). Broadening the ethnography to cover non-public interaction, and non-online interaction could reveal even more about lurkers' community involvement.

Lurking elsewhere

Synchronous vs. asynchronous: For the interviewees in the first study, lurking was confined to asynchronous environments. In these environments lurking is an invisible process, one in which the lurker can read posts, side post to other group members, or even decide not to participate in any fashion. In contrast synchronous environments do not generally lend themselves to lurking. There are a couple of reasons for this. When participants enter synchronous environments they usually do so to engage in conversation, and second, on entering the environment ones presence is generally visible to the other participants. If lurking occurs in synchronous environments, it will likely be very different from that of asynchronous lurking. Understanding that difference could be useful to designers of both types of environments.

DLs vs. other asynchronous environments: The second study focused on lurking in DLs as it would not have been possible to measure lurking levels using posting data from either newsgroups or BBSs. However, it is important to understand the limitations of focusing on DLs by examining some of the differences between DLs and both newsgroups and BBSs. Perhaps the most important difference is that DL messages are received as email. DL email competes with other types of email for the attention of the subscriber.

While it is true that most email clients are capable of filtering and depositing email in separate mail boxes, this has not been shown to be the practice of most email users (Nonnecke, 2000; Whittaker & Sidner, 1996).

In contrast to DLs accessed through email clients, Web-based BBSs and newsgroups are accessed through specially built user interfaces. This separates group communication from other non-group communication. Furthermore, the act of retrieving messages from either a newsgroup or a BBS is conscious and deliberate. Email clients often perform the task of retrieving e-mail automatically, e.g., once every 10 minutes. Email clients can also be used to get or check for email on demand. What is not known is whether an active vs. a passive process of obtaining messages has any impact on participation, e.g., reading, browsing, or posting.

There are two other major differences between DLs and the other tools. Firstly, email-based DLs poorly show conversational threading, and secondly, messages can be received as a digest (a single large email containing a set of messages for the purpose of reducing the volume of email). In both cases, the onus is on the receiver to reconstruct conversational threads. If the continuity of subject headings is to be maintained in the DL, replying to a message received in digest form requires the reply message's subject header to be manually constructed. The lack of visible threading and awkwardness of replying is being addressed by recent advances in digest-reader software (TECHWR-L, 1999), but it is not yet a common feature in email clients. In high traffic DLs, the lack of threading and digest format may make it harder to follow conversations. This in turn may make it more difficult to publicly join in the conversation. Given that this thesis has focused on DLs, understanding how lurking differs in other environments is an interesting follow-up to the current research.

Voluntary vs. mandatory learning: Work on vicarious leaning has been carried out by a number of researchers (Lee, McKendree, Dineen, & Mayes, 1999). The basic assumption is that students can learn by being exposed to interaction between students, and between students and teachers. In effect, the public discussion in learning groups are community goods that act as learning aids. The vicarious learner (i.e., a lurker) can find value in these goods and benefits not only themselves but the community as they do not draw further resources away from the learning community. An interesting follow-up to this thesis would be a comparison of mandated vicarious learning and the voluntary vicarious learning in health-support groups.

Joining vs. browsing: In the first study (Nonnecke & Preece, 1999; Nonnecke & Preece, 2000b), several participants described subscribing to a DL as a form of commitment with associated responsibilities to the other members. They also felt posting to a DL increased their commitment to the group and the presence created through posting should be maintained. Most DLs reinforce this by sending out a welcome message outlining what is expected of members in terms of participation and behaviour. By contrast, there is no subscription process for most BBSs and newsgroups. As a result, participation in DLs may differ from either BBSs or newsgroups, due to a different sense of responsibility to the group.

8.3 Conclusion

While work on this thesis was taking place, a question was put forward at a number of conferences and workshops on online groups and communities: does anyone know of any research on lurkers? Nobody came forward with a name of a paper or an author. However, many came forward with opinions, ideas, and personal experiences. Non-researchers have shown a similar interest in this work. The interviewees in the first study were very interested in having their opinions heard, and friends and family of the author have been similarly enthusiastic about sharing their lurking experiences. It is obvious that the topic has strong universal appeal, which no doubt comes from the fact that each and every one of us has either lurked, is lurking or will lurk in the future. As one researcher said, "lurking is normal, it is the people who post who are abnormal" (B. Wellman, personal communication, May 1999).

Summary results from the two studies are shown in Table 8.1. Based on the results from the interview study it is safe to say that lurking is widespread. All interviewees said they lurked and some lurked all the time. This was corroborated in the log study where more than 55% of the DL members lurked with no posts

(81% if lurking is defined as 1 or fewer posts/month). While this is lower than the oft quoted 90% figure (Mason, 1999), it represents a large number of participants.

A *primary* question in this thesis was *Why do lurkers lurk?* The unexpected outpouring of 117 reasons gives a sense of the complexity of lurking. This complexity is added to when one considers the many activities of lurking. The interviewees activities and rationales offer insight into all group members. Message selection, deleting, archiving and reading are activities common to all members, not just lurkers. Lurking is not the single simple action of not posting. Even public participants do not post all the time. In their moments of non-posting, they could be considered lurkers.

Contrary to what has been said elsewhere (Kollock & Smith, 1996), lurkers do not appear to be free-riders. Their non-public participation as lurkers is both beneficial and an acceptable part of online participation. Traditional definitions on participation emphasize public participation with very little understanding of non-public participation. Viewing online groups and communities through public participation only, casts lurkers in an unfavourable light, one they do not deserve (based on the research into lurkers carried out in this thesis.)

Lurking was found to be lower in health-support DLs, than in software-support DLs. This may be the result of a less technically sophisticated membership in health-support DLs, but may be also the result of the community type. As is already known, communication in many health-support groups is predominated by displays of empathy (Preece, 1998; Preece & Ghozati, 1998).

Three models of lurking (*filter*, *gratification* and *persistence*) were developed to account for lurkers' processes, needs and circumstances. These models provide insight into the context of lurking, shedding light on why lurkers lurk. The models present lurking as a situated activity in the context of life both inside and outside of online groups.

Lurkers are capable of having a sense of community. If judged by traditional definitions of community membership they do not meet all the requirements. However, much of their community mindedness and membership does not appear in the online public forum. This may be true for all members of online communities. That is, the online public interaction of the community may represent only a small portion of a community's total interaction. In any case, whether lurkers meet the definition of community members is a moot point, as they can and do feel a sense of community.

One of the most surprising findings of the log-based study was that DLs with high traffic rates also had the lowest lurker levels. This flies in the face of the interview results which suggested that lurking occurs when high traffic rates become burdensome. Why this is the case is currently unknown, but is certainly worth investigating. It may provide insight into creating and sustaining a thriving community.

A series of design implications were developed. These implications cover email clients, DL server software administration, supporting information for DL members, access mechanisms other than email clients, and changes at the member level. These design implications will be of value to DL members, moderators, designers, administrators and builders of online groups.

	Questions asked in this thesis	Finding
	Addressed in interview study	
P1	**Why do lurkers lurk?**	Lurkers have many and varied reasons for lurking. The *filter* model is put forward to organize and make sense of the reasons.
R1a	What motivates lurkers?	Lurkers have needs and lurking is a way of satisfying those needs. The *gratification* model describes the relationship between needs and lurking.
R1b	What role does lurking play in learning about the group?	Lurking is a means for getting to know a group. This is done through following threads, understanding individuals and many other ways.
R1c	How does persistent conversation affect lurking?	Persistent messages can help the lurker lurk, inhibit posting, and make work for the lurker. A *persistence* model of lurking is presented.
R1d	How do individual and group character differences affect lurking?	Both act as filters to posting. These elements are part of the *filter* model.
P2	**What do lurkers do?**	Interviewees employed five different strategies in their lurking.
R2a	What are the constraints on lurkers' activities?	Primary constraints are time and work.
	Addressed in logging study	
P3	**How many lurkers are there?**	Fewer than expected: still high with an average of over 55% for all DLs (when defined as 0 posts in 3 months).
R3a	Does lurking in health and software-support DLs differ?	Yes: health-support groups have lower levels of lurking (45% vs. 82%).
R3b	If lurking is defined as no posting, what happens to lurking levels when the definition is broadened to include minimal levels of posting, e.g., 1 post/month?	Lurking increases rapidly and then levels off as definition is broadened. Health-support groups maintain their lower levels of lurking (75% vs. 97% for software when lurking is defined as 3 or fewer posts/3 months).
R3c	Is there a relationship between lurking and the number of members in the DL?	Yes: smaller DLs have fewer lurkers.
R3d	Is there a relationship between lurking and the traffic level of the DL?	Yes: higher traffic means lower lurking.
R3e	If posting is concentrated with a few posters, how does that affect lurking levels?	The greater the concentration, the less the lurking.
R3f	Are short messages related to lower levels of lurking?	Yes: short messages are related to lower levels of lurking.
R3g	If clumpiness is an indication of interaction, does it necessarily follow that increased clump size is related to decreased lurking?	Yes: larger clumps are related to lower levels of lurking.
R3h	Is there a relationship between the number of singleton posters (and singleton posters who do not receive a response) and level of lurking?	Yes: as the number of singleton posters rises (and those who do not receive a response), so does the lurking.

Table 8.1: Summary of questions and findings from the two studies.

The definition of lurker provided by the Jargon Dictionary was reviewed and improved upon. The following definition captures the spirit of the term, *lurker*, as revealed in this thesis:

Extended definition of lurker: The term, *lurker*, is frequently used pejoratively and usually refers to anyone who never posts or posts infrequently. In fact, lurking is non-public participation. Lurking is a situated action, and many personal and group-, work-, and tool-related factors affect the activities and level of public and non-public participation. Lurking is "normal" in the sense that everyone is likely to be a lurker at some point in time. Lurkers are heterogeneous in most respects except in their lack of public posting. Therefore, in the absence of an understanding of the context in which it takes place, *lurker* is a meaningless term. Avoidance of the term, *lurker*, is recommended. Instead, the term *non-public participant (NPP)* is suggested. *NPP* is not pejorative and suggests there are other forms of valid participation outside of public posting.

References

Ackerman, M., & Starr, B. (1995). Social activity indicators: Interface components for CSCW systems. *Proc. Proceeding of the ACM Symposium on User Interface Software and Technology Con.*

Ackerman, M., & Starr, B. (1996). Social activity indicators for groupware. *IEEE Computer, 29*(6), 37-42.

ACOR. (1999). *About Us*, [Web page]. Available: http://www.acor.org/about/index.html.

Anderson, T., & Kanuka, H. (1997). On-line forums: New platforms for professional development and group collaboration. *Journal of Computer-Mediated Communication, 3*(3).

Barry, C. A. (1995). Critical issues in evaluating the impact of IT on information activity in academic research: Developing a qualitative research solution. *Library & Information Science Research, 17*, 107-134.

Barry, C. A. (1997). The research activity timeline: A qualitative tool for information research. *Library & Information Science Research, 19*(2), 153-179.

Beaudouin, V., & Velkovska, J. (1999, Jan. 22-24). The Cyberians: an empirical study of sociality in a virtual community. *Proc. Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities Con.*, Edinburgh.

Bellotti, V., & Rogers, Y. (1997). From Web press to Web pressure: Multimedia representations and multimedia publishing. *Proc. Conference on Human Factors in Computing Systems, CHI'97 Con.*

Bennahum, D. S. (1996). *The hot new medium is email* (Wired Collections/New Media/Virtual Communities), [Web page]. Wired. Available: http://www.hotwired.com/collections/virtual_communities/6.04_email1.html.

Berge, Z. (1992). *The role of moderator in Scholarly Discussion Groups (SDG).* Available: IPCT logs, archived at listserv@listserv.georgetown.edu.

Brand, S. (1994). *How Buildings Learn: What Happens After They're Built*. New York: Penguin Books.

Collins, M. P., & Berge, Z. L. (1997, March 24-28). Moderating online electronic discussion groups. *Proc. American Educational Research Association Annual Meeting Con.*, Chicago, IL.

Denzin, N. K., & Lincoln, Y. S. (1994). *Handbook of Qualitative Research*. London: Sage.

Dillon, A. (1998). Cultural analysis and what designers need to know - A case of sometimes too much, sometimes too little, and always too late. *ACM SIGDOC Journal of Computer Documentation, 22*(1), 12-17.

Donath, J., Karahalios, K., & Viegas, F. (1999). Visualizing Conversation. *Proc. HICSS-32 (Thirty-second Hawaii International Conference on System Science) Con.*, Maui, HI.

Donath, J. S. (1996). *Inhabiting the virtual city: The design of social environments for electronic communities.* Unpublished Ph.D., Massachusetts Institute of Technology, Cambridge, Mass.

Erickson, T. (1997). Social interaction on the net: Virtual community as participatory genre. *Proc. Thirtieth Hawaii International Conference on System Science (HICSS-30) Con.*, Maui, HI.

Erickson, T. (1999). *Call for Participation for a Workshop and Minitrack on "Persistent Conversation" (HICSS-33)*, [Web page]. Available: http://www.pliant.org/personal/Tom_Erickson/CFP.HICSS2000.html.

Excite.com. (1999). *What We Do*, [Web page]. Excite Network. Available: http://corp.excite.com/Company/what.html.

Fetterman, D. M. (1998). *Ethnography: Step by Step*. (2nd ed.). (Vol. 17). Thousand Oaks, California: Sage Publications, Inc.

Garton, L., Haythornthwaite, C., & Wellman, B. (1997). Studying online social networks. *Journal of Computer-Mediated Communication, 3*(1).

Grudin, J., & Palen, L. (1995). Why groupware succeeds: Discretion or mandate. *Proc. Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work Con.*

Gunnarsson, B. L. (1997). The writing process from a sociolinguistic viewpoint. *Written Communication, 14*(2), 139-188.

Harper, B. D., & Norman, K. L. (1993). Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. *Proc. Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference Con.*, Virginia Beach, VA.

Hendry, D., Nonnecke, B., Lungu, D., Carey, T., Mitterer, J., & Sobiesiak, R. (1990). *Writing with the BookManager program in mind: Examples, summary of use and writing advice* (IBM Technical Report TR 74.063). Toronto: IBM Canada Ltd.

Hewson, C., Laurent, D., & Vogel, C. (1996). Proper methodologies for psychological and sociological studies conducted via the Internet. *Behavior Research Methods, Instruments & Computers, 28*(2), 186-191.

Hiltz, R. (1993). *The Virtual Classroom: Learning Without Limits via Computer Networks.* Norwood, N.J.: Ablex Publishing.

Holtzblat, K., & Beyer, H. (1995). Contextual Design: using customer work models to drive system design. *Proc. CHI'95 Conference on Human Factors in Computing Systems Con.*

Hughes, J., King, V., Roden, T., & Andersen, H. (1994). Moving out from the control room: Ethnography in system design. *Proc. Computer Supported Cooperative Work Con.*, Chapel Hill, NC.

Jargon Dictionary. (1999). *Lurker Definition*, [Web page]. Available: http://www.netmeg.net/jargon.

Jones, Q. (1997). Virtual communities, virtual settlements & cyber-archaeology: A theoretical outline. *Journal of Computer-Mediated Communication, 3*(3).

Karetnick, R. D. (1998). A whole new world: sociolinguistic markers in computer-mediated multi-user communities. *Poster at NCA Con.*, New York.

Katz, J. (1998). *Luring the Lurkers*, [Web page]. Slashdot. Available: http://slashdot.org/features/98/12/28/1745252.shtml.

Kayany, J. M. (1998). Contexts of uninhibited online behavior: Flaming in social newsgroups on Usenet. *Journal of The American Society for Information Science, 49*(12), 1135-1141.

King, S. A. (1994). Analysis of electronic support groups for recovering addicts. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century, 2*(3), 47-56

Kollock, P., & Smith, M. (1996). Managing the virtual commons: cooperation and conflict in computer communities. In S. Herring (Ed.), *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives* (pp. 109-128). Amsterdam: John Benjamins.

Kollock, P., & Smith, M. (1999). Introduction: Communities in Cyberspace. In P. Kollock & M. Smith (Eds.), *Communities in Cyberspace*. London: Routledge.

Kraut, R. (1996). The Internet@Home. *Communications of the ACM, 39*(12), 32-35.

Kraut, R., & Atwell, P. (1997). Media use in a global corporation. In S. Kiesler (Ed.), *Culture on the Internet* (pp. 323-342). Mahway, NJ: Lawrence Erlbaum Associates.

Kraut, R., Mukhopadhyay, T., Szczypula, J., Kiesler, S., & Scherlis, B. (1998a). Communication and information: Alternative uses of the Internet in households. *Proc. CHI 98 Conference on Human Factors in Computing Systems Con.*, Los Angeles, CA.

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukopadhyay, T., & Scherlis, W. (1998b). Internet paradox: A social technology that reduces social involvement and psychological well-being. *American Psychologist, 52*(9), 1017-1031.

Kraut, R., Scherlis, W., Mukhopadhyay, T., Manning, J., & Kiesler, S. (1996). HomeNet: A field trial of residential Internet services. *Proc. CHI 96 Con.*, Vancouver, BC.

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

Lee, J., McKendree, J., Dineen, F., & Mayes, T. (1999). Learning vicariously in a distributed environment. *Active Learning, 10*(July), 4-9.

Lehnert, W. G. (1998). *Internet 101: A Beginners Guide to the Internet and the World Wide Web*. Reading, MA: -Wesley Longman Inc.

L-Soft International. (1999a). *List search*, [Web page]. Available: http://www.lsoft.com/lists/list_q.html.

L-Soft International. (1999b). *List with 1,000 subscribers or more*, [Web page]. Available: http://www.lsoft.com/scripts/wl.exe?XS=1000.

L-Soft International. (1999c). *Record sites*, [Web page]. Available: http://www.lsoft.com/listserv.stm.

Macaulay, C. (1999). Inscribing the palimpsest: Information sources in the newsroom. *Proc. Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities Con.*, Edinburgh.

Mason, B. (1999). Issues in virtual ethnography. *Proc. Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities Con.*, Edinburgh.

McGrath, J. E. (1992). Groups interacting with technology: The complex and dynamic fit of group, task, technology, and time. *Proc. CSCW'92 Conference on Computer-Supported Cooperative Work Con.*

McMillan, S. J. (1998). Who pays for content? Funding in interactive media. *Journal of Computer-Mediated Communication, 4*(1).

Merriam-Webster. (1999). *Lurk* (1999), [Web page]. WWWebster Dictionary. Available: http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=lurk.

Millen, D. R. (1997). Information technology and the computer-assisted reporter. *Proc. Summer CASTAC Conference (The committee on the Anthropology of Science, Technology, and Computing), Renneselaer Polytechnic Institute Con.*, Troy, NY.

Millen, D. R., & Dray, S. M. (1999). Information sharing in an online community of journalists. *Proc. Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities Con.*, Edinburgh.

Morris, M., & Ogan, C. (1996). The Internet as mass medium. *Journal of Computer-Mediated Communication, 1*(4).

Newhagen, J. E., & Rafaeli, S. (1996). Why communication researchers should study the Internet: A dialogue. *Journal of Computer-Mediated Communication, 4*(4).

Nonnecke, B. (1992). *Usage monitors as exploratory tools in the study of online documentation.* Unpublished M.Sc., University of Guelph, Guelph.

Nonnecke, B., & Preece, J. (1999, Jan. 22-24). Shedding light on lurkers in online communities. *Proc. Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities Con.*, Edinburgh.

Nonnecke, B., & Preece, J. (2000a). Lurker Demographics: Counting the Silent. *Proc. ACM CHI 2000 Con.*, The Hague.

Nonnecke, B., & Preece, J. (2000b). Persistence and lurkers: A pilot study. *Proc. HICSS-33 Con.*, Maui, Hawaii.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

Parks, R. M., & Floyd, K. (1996). Making friends in cyberspace. *Journal of Computer-Mediated Communication, 1*(4).

Preece, J. (1998). Empathic Communities: Balancing emotional and factual communication. *Interacting with Computers: The interdisciplinary Journal of Human-Computer Interaction, 5*(2), 32-43.

Preece, J. (2000). *Online Communities: Supporting Sociability and Designing Usability*. In Press.

Preece, J., & Ghozati, K. (1998). Empathy online: A review of 100 online communities. *Proc. Association for Information Systems, 1998 Americas Conference Con.*, Baltimore, MD.

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

Rafaeli, S., & Sudweeks, F. (1997). Networked interactivity. *Journal of Computer-Mediated Communication, 2*(4).

Roberts, T. L. (1998). Are newsgroups virtual communities? *Proc. CHI 98 Conference on Human Factors in Computing Systems Con.*, Los Angeles, CA.

Schiano, D. J., & White, S. (1998). The first noble truth of cyberspace: People are people (even when they MOO). *Proc. CHI 98 Con.*, Los Angeles, CA.

Shenk, D. (1997). *Data smog - Surviving the information glut*. New York: Harper Collins.

Slashdot. (1999). Available: http://slashdot.org.

Smith, C. B. (1997). Casting the net: Surveying an Internet population. *Journal of Computer-Mediated Communication, 3*(3).

Smith, M. (2000). Data mining social cyberspace. *Unpublished manuscript*.

Smith, M., Farnham, S. D., & Drucker, S. M. (2000). The social life of a small graphical chat spaces. *Proc. CHI 2000 Con.*, The Hague.

Smith, M. A. (1999). Invisible crowds in cyberspace: Mapping the social structure of the Usenet. In M. A. Smith & P. Kollock (Eds.), *Communities in Cyberspace* (pp. 195-219). London: Routledge.

Sproull, L., and Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science, 32*(11), 1492-1512.

Squires, D. (1997). Exploring the use of interactive information systems in academic research: borrowing from the ethnographic tradition. *Education for Information, 15*, 323-330.

Sternstein, M. (1996). *Statistics*. Hauppauge, NY: Barron's Educational Series, Inc.

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

Sudweeks, F., & Simoff, S. J. (1999). Complementary explorative data analysis: The reconciliation of quantitative and qualitative principles. In S. Jones (Ed.), *Doing Internet Research: Critical Issues and Methods for Examining the Net* (pp. 29-56). London: Sage Publications Ltd.

TECHWR-L. (1999). *About the Digest Reader*, [Web page]. David Castro. Available: http://www.raycomm.com/techwhirl/digestreader.html.

Thomsen, S. R., Straubhaar, J. D., & Bolyard, D. M. (1998). Ethnomethodology and the study of online communities: Exploring the cyber streets. *Proc. IRISS'98 Con.*, Bristol, UK.

Viegas, F. B., & Donath, J. S. (1999). Chat Circles. *Proc. CHI 99 Con.*, Pittsburgh, PA.

Walsham, G. (1993). *Interpreting Information Systems in Organizations*. Chichester: John Wiley & Sons.

Walther, J. B., Burgoon, J. K., & Park, D. W. (1994). Interpersonal effects in computer-mediated interaction: A meta-analysis of social and antisocial communication. *Communication Research, 19*, 52-90.

Weil, D. (1999). [Web page] Available: http://www.pathfinder.com/drweil/.

Wellman, B. (1997). An electronic group is virtually a social network. In S. Kiesler (Ed.), *Culture of the Internet* (pp. 179-208). Mahwah, NJ: Lawrence Erlbaum Associates.

Wellman, B., & Gulia, M. (1999). Virtual communities as communities: Net surfers don't ride alone. In M. A. Smith & P. Kollock (Eds.), *Communities in Cyberspace* (pp. 167-194). London: Routledge.

Whittaker, S., Isaacs, E., & O'Day, V. (1997). Widening the net. *SIGCHI Bulletin, 29*(3), 27-30.

Whittaker, S., & Sidner, C. (1996). Email overload: Exploring personal information management of email. *Proc. CHI'96 Con.*, Vancouver, BC.

Whittaker, S., Terveen, L., Hill, W., & Cherny, L. (1998). The dynamics of mass interaction. *Proc. CSCW 98 Con.*, Seattle, WA.

Wilder, L. (1999). *7 Steps to Fearless Public Speaking*. NY: John Wiley & Son, Inc.

Winzelberg, A. (1997). The analysis of an electronic support group for individuals with eating disorders. *Computers in Human Behavior, 13*(3), 393-407.

Worth, E. R., & Patrick, T. B. (1997). Do electronic mail discussion lists act as virtual colleagues. *Proc. AMIA Annual Fall Symposium Con.*

Appendices

Index of Tables

Appendix A: List of DLs

List name	Members	Traffic (posts/day)	No. lurkers with 0 posts	Lurkers with 0 posts (% of membership)	Lurkers with 3 or fewer posts (% of membership)
aamdstalk	272	14.89	103	37.87	72.06
addisons	88	15.50	13	14.77	59.09
adencyst	145	9.87	34	23.45	62.07
alpha1	312	40.26	63	20.19	65.06
amazon	150	22.81	68	45.33	73.33
apraxiakids	1079	29.82	610	56.53	83.60
assist	68	1.08	37	54.41	85.29
atfamilyl	118	4.82	47	39.83	72.03
backsl	238	0.62	211	88.66	98.32
behcet	212	8.07	109	51.42	81.13
bladderonc	67	23.94	0	0.00	2.99
braintmr	1345	72.29	545	40.52	75.24
cadis	588	30.18	363	61.73	80.61
carcinoid	223	28.20	0	0.00	46.19
caregivers	269	23.08	125	46.47	75.46
cdlskids	89	20.31	3	3.37	32.58
celiacdiabetes	148	0.70	130	87.84	95.95
celkids	446	0.12	439	98.43	100.00
cfsl	2297	0.38	2289	99.65	99.91
cll	1027	16.69	649	63.19	92.11
cocure	1014	3.45	961	94.77	98.72
colon	519	37.27	185	35.65	71.29
copd	676	63.20	145	21.45	68.20
ctclmf	334	10.79	132	39.52	82.93
deafl	363	23.21	43	11.85	78.51
desmoid	144	20.67	25	17.36	61.81
dialysispro	23	0.70	13	56.52	86.96
enuresis	99	0.40	83	83.84	97.98
epilepsyl	494	18.04	235	47.57	81.78
esarcoma	119	8.61	35	29.41	61.34
facingahead	154	44.69	0	0.00	45.45
fibroml	1220	97.52	612	50.16	75.57
gendiseasej	234	3.07	168	71.79	92.31
guaisupport	433	50.17	186	42.96	64.43
gynonc	171	12.27	97	56.73	74.85
headneckonc	75	5.63	3	4.00	68.00
hearttalkl	121	5.26	73	60.33	77.69
hemonc	848	4.35	713	84.08	96.82
hgfadults	62	2.36	25	40.32	80.65
hgfpeds	97	10.23	0	0.00	42.27

Table A1(part 1): List of health-support DLs included in demography

List name	Members	Traffic (posts/day)	No. lurkers with 0 posts	Lurkers with 0 posts (% of membership)	Lurkers with 3 or fewer posts (% of membership)
holisticmed	93	1.01	61	65.59	96.77
kidneyonc	300	33.29	15	5.00	53.00
larynxc	144	3.62	84	58.33	83.33
lca	310	3.61	231	74.52	91.61
lcdiabetes	386	11.18	244	63.21	84.46
lhon	59	0.45	48	81.36	94.92
liversupportl	245	19.64	77	31.43	67.76
lowcarblist	1952	57.04	1269	65.01	85.81
lungonc	281	13.67	94	33.45	74.02
lupus	404	16.99	283	70.05	85.40
lymphedema	246	11.79	93	37.80	73.17
mdlist	166	27.71	0	0.00	48.19
memoryprobs	15	0.24	8	53.33	93.33
mpdnet	1095	19.70	782	71.42	91.51
mpdsupportl	641	7.48	489	76.29	95.16
myeloma	811	35.13	386	47.60	78.79
nblastoma	103	28.62	0	0.00	27.18
osteop	79	0.76	50	63.29	94.94
ourkids	815	20.70	453	55.58	83.93
ovarian	789	40.23	465	58.94	81.24
pancreasonc	219	29.76	0	0.00	59.36
pedonc	222	26.67	47	21.17	60.81
pkusupportl	886	12.12	583	65.80	91.31
prostate	1442	24.43	1069	74.13	91.47
rblastoma	163	11.62	32	19.63	63.80
rossprocedure	81	3.26	18	22.22	74.07
rubber	357	23.56	155	43.42	71.71
sarcoma	165	28.17	32	19.39	60.00
sasyfras	279	34.65	77	27.60	60.93
sbparents	265	35.43	0	0.00	32.83
slfhlpl	80	0.02	78	97.50	100.00
solution	264	0.88	229	86.74	97.73
stomachonc	189	6.57	82	43.39	78.31
suicide	78	0.61	65	83.33	96.15
tcnetdig	385	29.60	178	46.23	72.47
thyroidonc	206	2.88	130	63.11	88.83
xxypeds	78	10.60	0	0.00	41.03

Table A1(part 2): List of health-support DLs included in demography

List name	Members	Traffic (posts/day)	No. lurkers with 0 posts	Lurkers with 0 posts (% of members)	Lurkers with 3 or fewer posts (% of members)
agiledogs	1971	45.98	1136	57.64	85.19
buslibl	1423	19.26	816	57.34	91.92
edil	1853	10.30	1525	82.30	96.92
equinel	1254	79.33	786	62.68	79.67
giftpl	1678	19.60	1183	70.50	92.97
gymnl	984	14.04	722	73.37	91.36
hlficl	1053	3.36	989	93.92	98.20
hteach	1582	5.85	1409	89.06	98.48
isogeochem	1247	3.49	1071	85.89	99.28
orthodox	1179	27.01	916	77.69	88.46
sheepl	1254	48.38	834	66.51	86.28

Table A2: List of large DLs included in demography

List name	Members	Traffic (posts/day)	No. lurkers with 0 posts	Lurkers with 0 posts (% of members)	Lurkers with 3 or fewer posts (% of members)
accmail	2566	15.02	2222	86.59	97.70
advancl	246	1.52	183	74.39	96.75
advhtml	4776	2.87	4654	97.45	99.71
deltal	337	0.30	323	95.85	99.70
facsupl	715	0.74	669	93.57	99.86
frontpage	255	3.85	166	65.10	87.45
genstat	233	0.96	191	81.97	97.85
lanmanl	672	9.13	360	53.57	93.75
listowners	70	0.21	57	81.43	100.00
machelp	149	2.27	83	55.70	90.60
macsystm	74	0.10	68	91.89	100.00
oraedusig	94	0.13	84	89.36	100.00
orausf	121	0.07	117	96.69	100.00
outlookl	274	2.76	194	70.80	94.89
pcask	443	6.36	325	73.36	91.65
pcsoft	1002	15.61	653	65.17	90.62
pctechl	159	0.48	139	87.42	98.74
procite	846	1.57	774	91.49	99.29
qpsusers	327	0.30	315	96.33	99.39
smagic	121	1.05	93	76.86	95.04
win97	430	0.18	421	97.91	99.77

Table A3: List of software-support DLs included in demography

Appendix B: Frequency tables

List type		Lurkers with 0 posts (% of membership)	Lurkers with 3 or fewer posts (% of membership)
health	**N**	**77**	**77**
	Mean	45.5565	75.1420
	Minimum	.00	2.99
	Maximum	99.65	100.00
	Std. Error of Mean	3.2688	2.2151
	Std. Deviation	28.6836	19.4375
large	**N**	**11**	**11**
	Mean	74.2641	91.7025
	Minimum	57.34	79.67
	Maximum	93.92	99.28
	Std. Error of Mean	3.8007	1.9111
	Std. Deviation	12.6056	6.3386
software	**N**	**21**	**21**
	Mean	82.0433	96.7983
	Minimum	53.57	87.45
	Maximum	97.91	100.00
	Std. Error of Mean	3.0410	.8468
	Std. Deviation	13.9357	3.8807
Total	**N**	**109**	**109**
	Mean	55.4832	80.9856
	Minimum	.00	2.99
	Maximum	99.65	100.00
	Std. Error of Mean	2.8300	1.8099
	Std. Deviation	29.5465	18.8956

Table B1: Frequency table for lurkers

List type		No. of members	Traffic (posts/day)	Message length (lines)	Clump size (messages)	No. of topics (or clumps)
health	**N**	**77**	**77**	**77**	**77**	**77**
	Mean	398.3636	18.4312	18.9749	2.5151	630.0779
	Minimum	15.00	.02	11.30	1.00	2.00
	Maximum	2297.00	97.52	93.59	4.54	3457.00
	Std. Error of Mean	50.1264	2.0940	1.1883	9.730E-02	75.5271
	Std. Deviation	439.8572	18.3750	10.4270	.8538	662.7473
large	**N**	**11**	**11**	**11**	**11**	**11**
	Mean	1407.0909	25.1439	32.9891	2.0418	973.0909
	Minimum	984.00	3.36	15.83	1.11	158.00
	Maximum	1971.00	79.33	138.07	2.71	3008.00
	Std. Error of Mean	97.7000	7.1600	10.7932	.1492	262.8351
	Std. Deviation	324.0341	23.7470	35.7970	.4949	871.7254
software	**N**	**21**	**21**	**21**	**21**	**21**
	Mean	662.3810	3.1179	15.7933	1.9643	118.9524
	Minimum	70.00	.07	6.27	1.00	3.00
	Maximum	4776.00	15.61	24.76	2.84	632.00
	Std. Error of Mean	238.1270	1.0149	1.1436	.1081	36.2277
	Std. Deviation	1091.2349	4.6510	5.2406	.4952	166.0161
Total	**N**	**109**	**109**	**109**	**109**	**109**
	Mean	551.0275	16.1584	19.7762	2.3612	566.2202
	Minimum	15.00	.02	6.27	1.00	2.00
	Maximum	4776.00	97.52	138.07	4.54	3457.00
	Std. Error of Mean	64.9670	1.7630	1.4257	7.657E-02	63.7673
	Std. Deviation	678.2751	18.4063	14.8852	.7994	665.7497

Table B2: Frequency table for membership, traffic, message length, and clumping

List type		Posters creating 1/4 of posts (% of posters)	Posters creating 1/4 of posts (% of members)
health	**N**	**77**	**77**
	Mean	5.050	1.987
	Minimum	.9	.0
	Maximum	50.0	6.7
	Std. Error of Mean	.753	.154
	Std. Deviation	6.608	1.347
large	**N**	**11**	**11**
	Mean	3.408	.849
	Minimum	1.2	.1
	Maximum	8.0	1.8
	Std. Error of Mean	.554	.154
	Std. Deviation	1.838	.512
software	**N**	**21**	**21**
	Mean	8.932	1.094
	Minimum	.3	.0
	Maximum	25.0	2.9
	Std. Error of Mean	1.404	.159
	Std. Deviation	6.436	.728
Total	**N**	**109**	**109**
	Mean	5.632	1.700
	Minimum	.3	.0
	Maximum	50.0	6.7
	Std. Error of Mean	.618	.121
	Std. Deviation	6.448	1.266

Table B3: Frequency table for posters creating ¼ of the posts

List type		Singleton posters (% of posts)	Singleton poster - no response (% of posts)
health	**N**	**77**	**77**
	Mean	36.7916	3.6389
	Minimum	.00	.00
	Maximum	100.00	100.00
	Std. Error of Mean	1.6316	1.3225
	Std. Deviation	14.3170	11.6046
large	**N**	**11**	**11**
	Mean	45.0827	4.9690
	Minimum	25.43	.69
	Maximum	63.07	10.24
	Std. Error of Mean	3.9830	1.0129
	Std. Deviation	13.2100	3.3594
software	**N**	**21**	**21**
	Mean	59.6010	11.5003
	Minimum	34.83	.00
	Maximum	90.00	63.64
	Std. Error of Mean	2.6714	3.0767
	Std. Deviation	12.2418	14.0993
Total	**N**	**109**	**109**
	Mean	42.0228	5.2877
	Minimum	.00	.00
	Maximum	100.00	100.00
	Std. Error of Mean	1.5701	1.1417
	Std. Deviation	16.3920	11.9196

Table B4: Frequency table for singleton posters

Appendix C: Independent sample test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
Variance assumption	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Lurkers with zero posts (% of members)									
equal	11.442	.001	-5.635	96	.000	-36.4869	6.4751	-49.3399	-23.6338
none			-8.172	68.760	.000	-36.4869	4.4646	-45.3941	-27.5796
Clump size (messages)									
equal	3.680	.058	2.713	96	.008	.1552	5.722E-02	4.165E-02	.2688
not			3.274	43.711	.002	.1552	4.741E-02	5.966E-02	.2508
Singleton posters (% of posts)									
equal	1.030	.313	-6.263	96	.000	-1.5521	.2478	-2.0440	-1.0602
not			-7.665	45.013	.000	-1.5521	.2025	-1.9600	-1.1443
Singleton posters - no response (% of posts)									
equal	1.518	.221	2.753	96	.007	5.893E-02	2.140E-02	1.644E-02	.1014
not			3.167	39.743	.003	5.893E-02	1.861E-02	2.131E-02	.09654
Traffic (posts/day)									
equal	3.629	.060	5.131	96	.000	1.4357	.2798	.8803	1.9911
not			6.108	42.487	.000	1.4357	.2351	.9615	1.9099
Posters creating1/4 of posts (% posters)									
equal	1.393	.241	-5.905	92	.000	-1.6994	.2878	-2.2709	-1.1278
not			-6.729	36.904	.000	-1.6994	.2526	-2.2111	-1.1876
No. of members									
equal	.099	.754	-1.169	96	.245	-.2962	.2535	-.7994	.2069
not			-1.109	29.716	.276	-.2962	.2672	-.8421	.2496

Table C: Independent sample test (health-support vs. software support)

Appendix D: Pearson correlation

	Lurkers with zero posts (% of members)	No. of lurkers with 0 posts	No. of members	Traffic (posts/day)	Mean message length (lines)
Lurkers with zero posts (% of members)		0.761 0.000	0.332 0.000	**-0.621** 0.000	0.258 0.007
No. of lurkers with 0 posts	0.761 0.000	.	0.760 0.000		0.195 0.043
No. of members	0.332 0.000	0.760 0.000	.	0.410 0.000	
Traffic (posts per day)	**-0.621** 0.000		0.410 0.000	.	
Mean message length (lines)	0.258 0.007	0.195 0.043		**-0.172** 0.074	.
Mean clump size	**-0.631** 0.000	**-0.489** 0.000	**-0.269** 0.005	0.421 0.000	**-0.344** 0.000
Number of topics (clumps)	**-0.442** 0.000		0.514 0.000	0.940 0.000	
Singleton posters (% of posts)	0.719 0.000	0.493 0.000		**-0.663** 0.000	0.177 0.066
Singleton poster - no response (% of posts)	0.723 0.000	0.469 0.000	0.219 0.025	**-0.667** 0.000	0.244 0.012
Posters creating ¼ of posts (% posters)	0.448 0.000		**-0.422** 0.000	**-0.777** 0.000	
Posters creating ¼ of posts (% membership)	**-0.662** 0.000	**-0.784** 0.000	**-0.739** 0.000		**-0.273** 0.004
Normalizing treatment	none	ln(x+1)	ln(x)	ln(x+1)	ln(x)

Note: Degree of association (R) is the upper, larger number. The lower number of the pair is the P value.

Table D1 (part 1): Pearson Correlation table

	Mean clump size	Number of topics (clumps)	Singleton posters (% of posts)	Singleton poster - no response (% of posts)	Posters creating ¼ of posts (% of posters)	Posters creating ¼ of posts (% of members)
Lurkers with zero posts (% of members)	**-0.631** 0.000	**-0.442** 0.000	0.719 0.000	0.723 0.000	0.448 0.000	**-0.662** 0.000
No. of lurkers with 0 posts	**-0.489** 0.000		0.493 0.000	0.469 0.000		**-0.784** 0.000
No. of members	**-0.269** 0.005	0.514 0.000		0.219 0.025	**-0.422** 0.000	**-0.739** 0.000
Traffic (posts/day)	0.421 0.000	0.940 0.000	**-0.663** 0.000	**-0.667** 0.000	**-0.777** 0.000	
Mean message length (lines)	**-0.344** 0.000		0.177 0.066	0.244 0.012		**-0.273** 0.004
Mean clump size	.	0.170 0.077	**-0.613** 0.000	**-0.762** 0.000	**-0.302** 0.001	0.472 0.000
Number of topics (clumps)	0.170 0.077	.	**-0.521** 0.000	**-0.492** 0.000	**-0.700** 0.000	
Singleton posters (% of posts)	**-0.613** 0.000	**-0.521** 0.000	.	0.835 0.000	0.436 0.000	**-0.450** 0.000
Singleton poster - no response (% of posts)	**-0.762** 0.000	**-0.492** 0.000	0.835 0.000	.	0.467 0.000	**-0.461** 0.000
Posters creating ¼ of posts (% of posters)	**-0.302** 0.001	**-0.700** 0.000	0.436 0.000	0.467 0.000	.	0.214 0.026
Posters creating ¼ of posts (% of members)	0.472 0.000		**-0.450** 0.000	**-0.461** 0.000	0.214 0.026	.
Normalizing treatment	ln(x+1)	sqrt(x)	sqrt(x+10)*	ln(x)	ln(x)	ln(x+1)

Note: Degree of association (R) is the upper, larger number. The lower number of the pair is the P value.

Table D1 (part 2): Pearson Correlation table

Appendix E: CHI 2000 paper

Appendix E contains the following refereed conference paper:

Nonnecke, B., & Preece, J. (2000). Lurker Demographics: Counting the Silent. *Proc. ACM CHI 2000 Con.*, The Hague.

Lurker demographics: Counting the silent

Blair Nonnecke
Dept. of Information Systems
1000 Hilltop Circle
Baltimore, MD 21250 USA
+1 410 455 3795
nonnecke@umbc.edu

Jenny Preece
Dept. of Information Systems
1000 Hilltop Circle
Baltimore, MD 21250 USA
+1 410 455 6238
preece@umbc.edu

ABSTRACT
As online groups grow in number and type, understanding lurking is becoming increasingly important. Recent reports indicate that lurkers make up over 90% of online groups, yet little is known about them.

This paper presents a demographic study of lurking in email-based discussion lists (DLs) with an emphasis on health and software-support DLs. Four primary questions are examined. One, how prevalent is lurking, and do health and software-support DLs differ? Two, how do lurking levels vary as the definition is broadened from zero posts in 12 weeks to 3 or fewer posts in 12 weeks? Three, is there a relationship between lurking and the size of the DL, and four, is there a relationship between lurking and traffic level?

When lurking is defined as no posts, the mean lurking level for all DLs is lower than the reported 90%. Health-support DLs have on average significantly fewer lurkers (46%) than software-support DLs (82%). Lurking varies widely ranging from 0 to 99%. The relationships between lurking, group size and traffic are also examined.

Keywords
Lurker, lurking, discussion list, demographic, newsgroup, BBS, email, health-support, traffic, membership

INTRODUCTION
DLs, newsgroups, and Web-based bulletin board systems (BBSs) have experienced rapid growth as the number of Internet users climbs. As of July 1999, there are more than 131,000 DLs using Listserv's® server software. The 69,000,000 members of these DLs receive in excess of 29,000,000 messages per day [6]. Whittaker et al [19] also cite large numbers for Usenet newsgroups. The growth and prevalence of online groups, coupled with the relative ease of gathering persistent and traceable messages, has made online groups a fertile ground for research. The following are a few of the areas so far studied: the development of friendship [12], the perception and quality of community [15], factors affecting interaction within newsgroups [19], and the development of empathy in health-support groups [13, 14]. Each of these studies was based on examining individuals participating in public spaces, i.e., those who post. None examined their chosen area from a lurking perspective, even though lurkers are reported to make up over 90% of several online groups [2, 7].

Given that lurkers are both unstudied and apparently in the majority, knowing more about them will have benefits in many areas. Their sheer number suggests they are important to study. From a usability perspective, improvements in tools and group design will fall out of a better understanding of lurkers and their activities. For lurkers and their communities, self-knowledge of lurking will demystify lurkers' roles, value, and activities. This has already been shown to be the case when a participant in an initial study responded to a draft article on lurking [11]:

> Maybe it's a sign of my own mild discomfort around being a lurker, but I found it reassuring to recognize myself and my behaviour within the continuum you describe, and to see lurking treated seriously, with both acceptance and respect. As a lurker, I'm used to observing from the sidelines and participating vicariously, and it's strangely gratifying to read an article that speaks directly to that experience. It's almost like suddenly feeling part of an (until-now) invisible community of lurkers.

In their pioneering work, Kollock and Smith [3] describe lurkers as free-riders, i.e., noncontributing, resource-taking members. Knowing more about lurkers and their lurking will show whether this is a fitting description.

As group and community development becomes an important component of commerce on the Internet, understanding lurkers will become an essential part of doing business. Every lurker is a potential

customer. For example, Amazon.com has been very successful in creating an online retail environment in which lurkers can make purchasing decisions based on how others have purchased in the past and on reviews supplied by other customers. Amazon.com has leveraged the information gained from those willing to post reviews into purchasing-support tools for the lurker and poster alike.

Definition

The online Jargon Dictionary [1] defines lurker as:

> One of the 'silent majority' in a electronic forum; one who posts occasionally or not at all but is known to read the group's postings regularly. This term is not pejorative and indeed is casually used reflexively: "Oh, I'm just lurking." When a lurker speaks up for the first time, this is called 'delurking'.

This definition suggests that lurking is the normal behaviour of the majority of the population and that lurking can be defined in terms of the level of participation, either as no posting at all or as some minimal level of posting. However, defining lurking is problematic. Should someone who never posts in public spaces, but regularly side-posts to individual group members, be deemed a lurker? If a person posts once and then never again, are they lurking? Is someone lurking when they go on holidays? Is someone lurking when for a period of time they do not post? While these are important considerations, this study takes the simple approach of defining lurking as either no posts or some minimal number of posts over a period of time.

Research questions

The work reported here is the second in a series of studies on lurking [10]. In the first study [11], Internet users were chosen for their membership in online groups, and not for their posting frequency. Given that lurking has been reported as a common means of participation [2, 7], it was assumed that lurkers and their behaviors would be readily encountered within the general Internet population. In the first study, it was found that each participant lurked in at least one online group, and several lurked in all of their online groups. This finding, among others, reinforced the need to better understand lurking. A demographic survey of online discussion groups would provide a different perspective from the first study by emphasizing quantitative measures.

DLs, rather than BBSs or newsgroups, were chosen as the basis of this study for a number of reasons. For the results to have their greatest value, the chosen communication technology needed to be widely used. L-Soft's usage figures show very high levels of use, and of the online discussion groups mentioned by participants in the first study, 25 of the 41 were DLs accessed through email. Just as importantly, DL servers track membership through their subscription mechanism. In turn, DL membership information can be accessed by querying a DL's server. The level of lurking can be measured by tracking posted messages and identifying posters. In contrast, membership levels are unavailable for most BBSs and newsgroups.

This study is an extension of work on online health-support communities [13, 14]. As well, in a study of who pays for content and interactive media, McMillan [8] provides several reasons for studying health-related groups:

> …health and health related subjects have in the past played a central role in the early financial support in many media; health related sites are the fastest growing topic areas in CMC; health-related sites are heavily used; and this area contains one of the fastest growing categories of consumer advertising.

For these reasons, health-support DLs are the focus of this investigation. For comparison purposes, software-support DLs are also included in this study.

The remainder of this paper examines four main questions:

Q1. How prevalent is lurking, and do health and software-support DLs differ?

Q2. If lurking is defined as no posting, what happens to lurking levels when the definition is broadened to include minimal levels of posting, e.g., 1 post/month?

Q3. Is there a relationship between lurking and the number of members in the DL?

Q4. Is there a relationship between lurking and the traffic level of the DL?

METHOD

The primary aim of this work is to understand how much lurking occurs in DLs, with specific emphasis on health and software-support groups.

Selection of DLs

To select DLs for the study, L-Soft's CataList catalog and DL search facility [4] were used to locate suitable DLs. A search on the word "support" resulted in a listing of 300 DLs and a description of each. From this listing, subscriptions were taken out on all public DLs relating to health or software-support. To increase sample size, additional subscriptions were taken out on a random selection of health (22) and software (10) support DLs. Although the additional DLs provide support for their members, neither their title nor their one-line catalog description contained the term "support".

DL set	1. Lurking (% of membership)			2. No. of members			3. Traffic (posts/day)		
	Mean	SD	SE	Mean	SD	SE	Mean	SD	SE
All N=109*	55.5	29.6	2.8	551	678.3	65.0	16.2	18.4	1.8
Health N=77*	45.5	28.7	3.3	398.4	439.9	50.1	18.4	18.4	2.1
Software N=21*	82.0	13.9	3.0	662.4	1091.2	238.1	3.1	4.7	1.0

* No. of DLs in set, SD=Standard Deviation, SE=Standard Error of Mean

Table 1: Lurking, no. of members and traffic for the DL sets

(Note: Analysis comparing these additional support DLs shows their lurking levels are not significantly different from those found through searching on "support", and as a result, they are included in this study.)

In addition to DLs related to health and software, a random set of DLs on other topics were selected for their large size (CataList displays a description of all DLs with membership greater than 1000 [5]). Eleven randomly selected DLs between 1000 and 2000 members were included as a basis for examining whether large DLs have a greater proportion of lurkers than smaller ones (see Q3. above).

Data collection

Messages were collected from the selected DLs over a three-month period at a rate of slightly less than 2000 messages/day. Eudora Pro was used to collect and filter email into separate mailboxes for each list, and to monitor the process on a regular basis. Using CataList, the membership size of each list was determined at the beginning and end of the collection period. Lurking levels reported in this study are based on the lower of the two membership levels recorded for the 12-week period.

At the end of 12 weeks, the DLs were examined to ensure that each DL had sent at least one post a month for the 12 weeks. Of the 135 original subscriptions, 109 DLs are included in the study. DLs were dropped from the study if they stopped sending messages for any reason, e.g., change of server, failure on the part of the researchers to reply to subscription notices, or a non-active DL. Messages from the remaining DLs were then run through a Perl script producing records containing the following fields: list name, date, time, size of message, subject heading, and sender. 147,946 messages were transcribed into records and imported into an SQL database. This provided an effective and flexible means for querying and analyzing the data. The data collected represents over 60,000 members and 19,000 posters.

RESULTS

Lurking levels

Q1. How prevalent is lurking, and do health and software-support DLs differ?

Using information from the SQL database, mean lurking levels were calculated for the set of all DLs, and for each of the health and software DL sets (see Table 1, column 1). Lurking was defined as no posts within the 12-week collection period. The mean lurking level for all DLs is less than the 90% figures reported by Katz and Mason [2, 7]. It should be noted that while the mean was less than 90%, 12% of the DLs had lurking levels higher than 90%.

The differences in mean lurking levels between the health and software-support DLs is significant. Software-support groups had almost double the number of lurkers. Figure 1 shows the distribution of the lurking levels for each DL type using a box and whisker display. (Note: See Sternstein [16, p. 37] for further information on this visual representation.) Each horizontal line represents a boundary for 25% of the DLs in the sample. The thicker line is also the median for each type. Each of the central boxes contains 50% of the DLs.

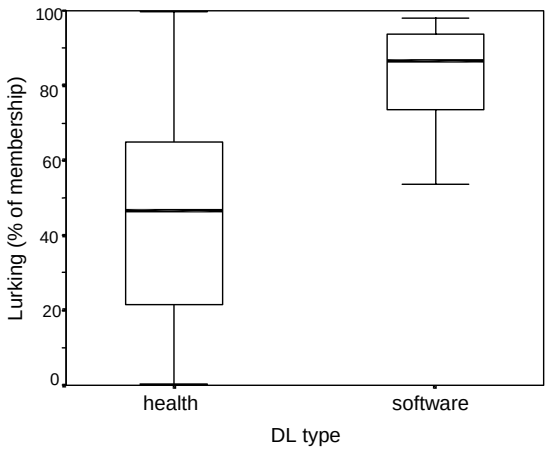

Figure 1: Distribution of lurking levels by quartile for each DL set

Software-support DLs show less variation and none have a lurking level of less than 50%. By contrast, the lurking level of health-support DLs range from zero to 99%. Health-support is a broad umbrella under which to investigate group behaviour. As such, lurking levels may vary according to a number of other factors, e.g., list topic, illness vs. injury, or chronic vs. short term disorders. This difference in variation between the two DL types may be the result of the greater number of health-support DLs in the study, which represents a broader cross-section of their type.

Apart from the low mean number of lurkers in the health-support DLs, what appears most striking about these results is the large variation in lurking levels, and that on average the lurking level for all DLs is lower than the reported 90% figure [2, 7].

Broadening the definition of lurking

Q2. If lurking is defined as no posting, what happens to lurking levels when the definition is broadened to include minimal levels of posting, e.g., 1 post/month?

In Table 1 lurking was defined as no posts during the 12-week collection period. If lurking is examined on a sliding scale where the allowable posting level can grow, a somewhat different picture emerges. In Figure 2, lurking levels were calculated for a range of cumulative posts, from no posts to 3 or fewer posts for the 12-week period (i.e., 1 or fewer posts per month). As the definition broadens to include more posts in the 12-week period (towards the 3 level), lurking levels move higher. At the level of 3 or fewer posts per 12-week period, the mean lurking level for the health DLs is still lower than 90%, while the software DLs' mean has moved above this level. Both the health and software-support DLs behave in a similar manner, and their relative offset is maintained.

A posting rate of 3 posts in 12 weeks is still an infrequent level of posting. It could be argued that most of what is being done by members at this level is not posting. Presence or visibility of members within a list may be a better indicator of lurking, i.e., is a member known to the other group members in a way that makes them somehow recognizable and thus not lurkers. Defining lurking as a function of the visibility of the poster suggests that other factors would influence this visibility, e.g., the number of members, the number of posters, the activity of the list, and the value and/or notoriety of each participant. It is possible that someone who flames on an irregular basis may be seen as less of a lurker than someone who contributes in a regular but less visible manner. The polar opposite of lurking may be stardom.

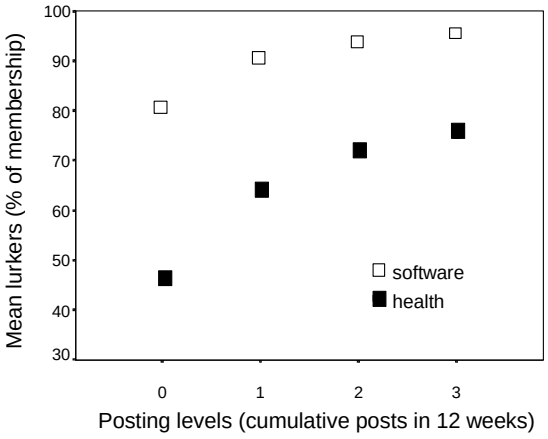

Figure 2: Variation of lurking levels for a range of cumulative posts.

Further work is needed in understanding lurking. For example, lurking may not be a continuous state and could be punctuated by periods of public posting based on topic or need. Using the current data set, there is no reason why analyses of this type cannot be carried out in the future. The raw data could also be used from a contextual or ethnographic perspective, one in which content and dialogue analyses could be carried out. Examples of these kinds of analyses can be found in Preece and Ghozati [14], and Worth and Patrick [20].

Lurking and the number of members

Q3. Is there a relationship between lurking and the number of members in the DL?

In large DLs lurking may be easier. As the number of members increases, the need for any given member to participate may decline. In addition, high posting levels could create chaos and lurking in large DLs may be a practical means of reducing the number of posts and maintaining order. If either of these is the case, then large DLs should have a greater proportion of lurkers than smaller ones. As can be seen in Table 1 (column 2) health-support DLs have on average fewer members than the software-support DLs. If increasing membership size has the effect of generating more lurkers, then the difference in mean membership levels could explain why health-support DLs have lower levels of lurking.

On examining all 109 DLs in the sample, the anticipated greater incidence of lurking in larger DLs is not strongly shown. Figure 3 shows a strong positive non-linear relationship between the number of lurkers and the size of the DL. A linear regression also fits this data equally well. If this result is taken at face value, membership size does not explain the

differences in lurking between the health and software DLs.

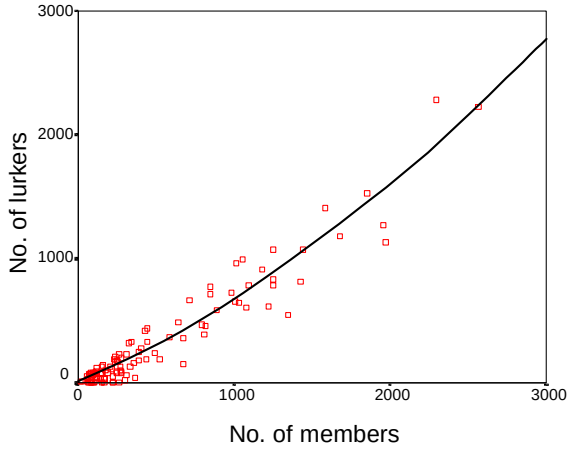

Figure 3: No. of lurkers vs. no. of members for each DL.

The relatively few DLs with over 500 members skews the relationship in favour of the larger DLs. Of the 98 health and software DLs, 74 of them have fewer than 500 members. Figure 4 is a scatter plot of these smaller DLs. The regression line in Figure 4 is a strong positive relationship with a slope less than that in Figure 3. This suggests that for DLs with fewer than 500 members, there are on average fewer lurkers than in the larger DLs. It should also be noted that that the software-support DLs in Figure 4 are distributed in a straight line. This suggests that even when software-support DLs are of equivalent membership size, they will on average have higher lurker levels.

DL members receive no direct information about the number of members in a DL. The cues that do exist are indirect, e.g., a query to the server for information, the number of different members posting, the variety of topics covered, and the traffic on the DL. It is possible that a DL of several thousand members could behave like and be indistinguishable from one with only 100 members. More work is required to understand how the size of DLs is perceived by members, and how members respond to this in their various forms of participation.

From the perspective of personal email management, once message rates get above a comfortable level, participating in a DL may take more effort, i.e., there are more messages to read, skim, reply to, etc. Based on participant input from the first study [9], traffic levels were divided into four categories requiring varying levels of management effort (see Table 2).

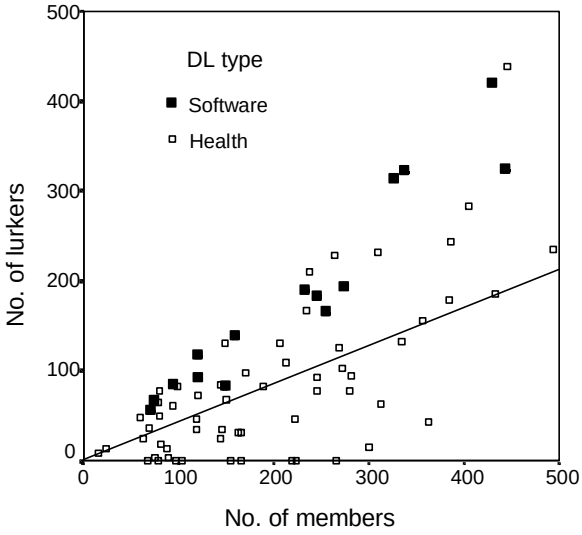

Figure 4: No. of lurkers vs. no. of members for each DL with less than 500 members

Lurking and DL traffic levels

Q4. Is there a relationship between lurking and the traffic level of the DL?

	Traffic level	
Management effort	messages/week	messages/day
None	< 1	< 0.14
Low	1-3	0.14-0.5
Medium	4-42	0.5-6.0
High	> 42	>6.0

Table 2: Traffic levels for a DL and the corresponding management efforts

The categorization was done prior to examining the distribution of posting rates from the current study. Several participants in the first study indicated that lists with less than 4 or 5 posts per day were easy to handle. In the current study, more than 50% of the DLs fall in the High category. It should also be noted that what is manageable will vary widely between individuals, and will depend on many factors, including type of email software, experience, demands on time, and interest.

Lurking levels for all DLs were negatively correlated with traffic (Pearson's correlation coefficient of -.426 is significant at the .01 level). Figure 5 shows that for a given DL size, lists with highest traffic levels generally have the lowest lurking levels. Banding by traffic level is visible, starting with the lowest traffic level (None) in the top left hand corner and progressing towards the lower lurking levels and larger DL size. This partially explains the lower

levels of lurking in health-support DLs as these had the higher traffic levels (see Table 1, column 3).

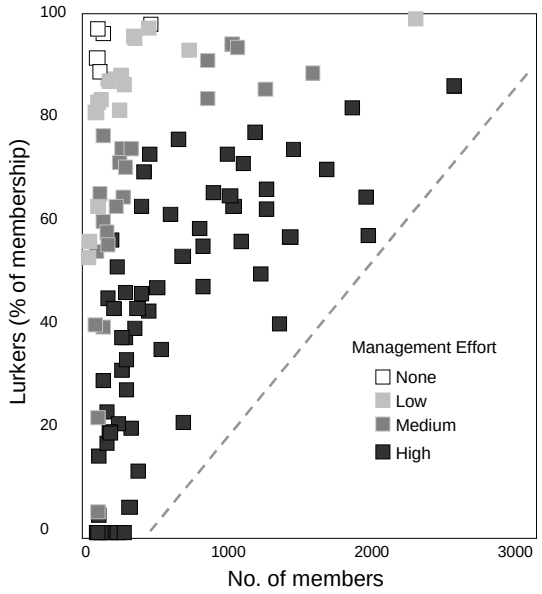

Figure 5: Lurker levels as related to management effort and number of members for each DL

Conspicuously absent are DLs in the area below the broken line, which appears to be a kind of interactive no-man's land. Why this should be the case is not known at the present time, but it could be related to the difficulty of making sense out of large DLs with high traffic volumes and large membership levels. At some point, the DL may become unusable and self-adjust through membership attrition and/or a decrease in public posting. It may be that lurking increases under conditions where having a public voice is difficult. In our initial study [11], several participants indicated they knew other people would post opinions similar to their own in active lists, and thus felt no need to post. When traffic is high, there is a sense that adding messages to the list only increases the traffic without improving the quality. For them, lurking was a way of reducing the noise on the list, a civic duty so to speak. It would be interesting to examine DLs that fall near or below the broken line, and determine whether they transform in any way, e.g., split, have high membership turnover, etc.

Below the 500-member level, health-support DLs appear evenly distributed with respect to number lurkers and thus lurking levels (see Figure 4). For these smaller, more personal-sized groups, the size of the DL may be less of an indicator of lurking level and some other factors may be at work. For DLs with fewer than 500 members, traffic levels appear to be a good predictor of lurking levels (see Figure 5). What drives the combination of low lurking levels and high traffic is still unclear, but may be related to the topic of the DL, motivation of members, and style of interaction (e.g., empathy vs. information exchange).

The DLs with high traffic levels are an interesting group (see Figure 6). The 11 DLs with average traffic levels over 40 messages/day had a low average lurking level of 44%. Four of the DLs were from the Large set of DLs and 7 were health-support. The median membership size for this group was high, at 1220. However, three of these high traffic DLs had fewer than 500 members. For the DLs in this high traffic range, lurking levels appear to be randomly distributed across membership size. As a result, high traffic levels don't appear to be a very good indicator of group size. It is possible that group size becomes immaterial to public participation when it isn't readily knowable.

DISCUSSION

Much of the discussion related to the four original questions can be found in the previous section. Therefore, this next section focuses on three important issues: lurkers as free-riders, traffic levels, and lurking elsewhere, i.e., how lurking in DLs may differ from either BBSs or newsgroups.

Lurkers as free-riders

In the Introduction it was mentioned that Kollock and Smith [3] describe lurkers as free-riders. Describing lurkers as free-riders classifies them for their lack of public participation and their use of resources without giving back to the group. Even when lurking is narrowly defined, e.g., less than one post/month, the vast majority of DL members are lurkers. This being the case, how do online groups survive in the face of almost universal free-riding?

One explanation is that lurking is not free-riding, but a form of participation that is both acceptable and beneficial to online groups. Public posting is but one way in which an online group can benefit from its members. Members of a group are part of a large social milieu, and value derived from belonging to a group may have far-reaching consequences, e.g., virus alerts being distributed beyond the posters of a DL specializing in combating viruses. A second explanation is that a resource-constrained model may not apply to online groups where the centralized cost of servicing 100 members isn't much different from that of serving 1000, or even 10,000. In large DLs the danger could be in not having enough lurkers.

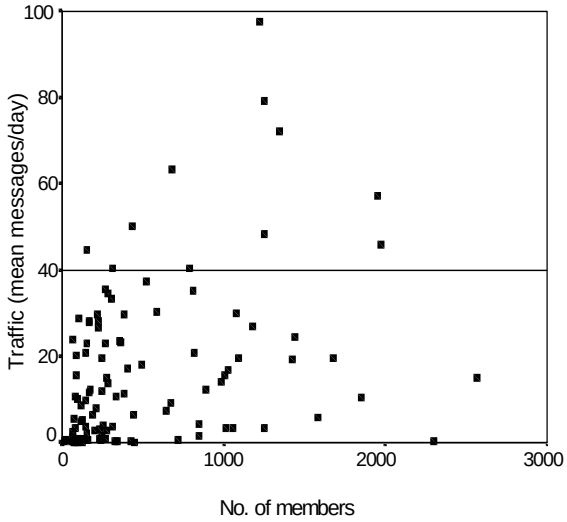

Figure 6: Traffic levels vs. no. of members

Traffic levels

In our first study [11], participants described the effort required to manage DL traffic. If there were few messages, then the DL was effectively out of mind and required little or no effort. If there were many messages, then the DL became burdensome. Several participants cited newsgroups as being less useful because of the large volume of messages. They also mentioned the quality of the messages as being very important, e.g., content, knowledge base of participants, and courtesy. Several participants left newsgroups because of what they perceived as low content quality.

Using figures supplied by participants, DLs with traffic levels of over 6 messages/day were categorized as requiring higher effort to manage. The mean traffic level for the software-support DLs was 3.1 posts/day (see Table 1, column 3). These values fit nicely with our expectations of manageable traffic. However, the mean traffic level of the health-support DLs was 18.4 messages/day, and one DL exceeded an average of 97 posts/day. These higher-than-expected numbers suggest that these DLs are somehow different than the DLs participants described as being ideal in the first study. Why the discrepancy? It is possible that these DLs supply such high-quality content that their members are willing to put in the higher effort to deal with them. It is also possible that high traffic DLs act like many little DLs, each identifiable by a set of subjects and/or authors. The observed high traffic levels suggest that what is an acceptable and perhaps necessary traffic level in one DL may be unacceptable in another. It also suggests that motivation, in addition to quality of messages, is an important facet of acceptable traffic levels.

Understanding what constitutes acceptable traffic rates is an important issues in designing online communities. E-commerce is already running into this problem. For example, when sending promotional materials through distribution lists, it is important to understand how much email can be sent before customers perceive it as a nuisance. Understanding how DL members cope with and make use of high volumes of messages is important for the designers of email-client software. Lastly, messages from DLs are not received in a vacuum; they compete with messages from a variety of other sources, including personal and professional correspondence, and email from other DLs -based email.

Lurking elsewhere

This study focused on lurking in DLs as it would not have been possible to measure lurking levels using posting data from either newsgroups or BBSs. However, it is important to understand the limitations of focusing on DLs by examining some of the differences between DLs and both newsgroups and BBSs.

Perhaps the most important difference is that DL messages are received as email. DL email competes with other types of email for the attention of the subscriber. While it is true that most email clients are capable of filtering and depositing email in separate mail boxes, this has not been shown to be the practice of most email users [9, 18].

In contrast to DLs accessed through email clients, Web-based BBSs and newsgroups are accessed through specially built user interfaces. This separates group communication from other non-group communication. Furthermore, the act of retrieving messages from either a newsgroup or a BBS is conscious and deliberate. Email clients often perform the task of retrieving e-mail automatically, e.g., once every 10 minutes. Email clients can also be used to get or check for email on demand. What is not known is whether an active vs. a passive process of obtaining messages has any impact on participation, e.g., reading, browsing, or posting.

There are two other major differences between DLs and the other tools. Firstly, email-based DLs poorly show conversational threading, and secondly, messages can be received as a digest (a single large email containing a set of messages for the purpose of reducing the volume of email). In both cases, the onus is on the receiver to reconstruct conversational threads. If the continuity of subject headings is to be maintained in the DL, replying to a message received in digest form requires the reply message's subject header to be manually constructed. The lack of visible threading and awkwardness of replying is being addressed by recent advances in digest-reader

software [17], but it is not yet a common feature in email clients. In high traffic DLs, the lack of threading and digest format may make it harder to follow conversations. This in turn may make it more difficult to publicly join in the conversation.

In our first study [11], several participants described subscribing to a DL as a form of commitment with associated responsibilities to the other members. They also felt posting to a DL increased their commitment to the group and the presence created through posting should be maintained. Most DLs reinforce this by sending out a welcome message outlining what is expected of members in terms of participation and behaviour. By contrast, there is no subscription process for most BBSs and newsgroups. As a result, participation in DLs may differ from either BBSs or newsgroups, due to a different sense of responsibility to the group.

The effects of different types of email tools and skills have been ignored in this study. However, this could be an important difference between health and software-support DLs and their participants. Software skill and acumen may vary for participants in these DL types. For example, members of software-related DLs may have better computer skills and a greater knowledge of the Internet than those of other DL types.

Personal characteristics that may impact lurking include motivation and comfort in communicating online. To investigate these other approaches are called for. These include member surveys and the examination of DLs from a content and dialogue perspective.

CONCLUSION AND FUTURE WORK

As this study shows, lurkers are everywhere, and that is OK. A case can be made for lurking being normal and public posting being abnormal. After all, if everyone were posting, who would be reading. It is unfortunate that the term lurker, with all of its negative connotation, has gained acceptance. Fortunately, lurking can now be understood as the many activities related to membership in online groups. Rather than being free-riders, lurkers should be called participants (publicly silent though they may often be).

As a quantitative follow-up of our interview-based study [11], this work proved a capable tool for understanding lurking. There is some irony in studying lurking with a method normally reserved for examining public participation. This work was successful in discovering a number of relationships between lurking levels, DL type, membership levels and traffic. Whether they are causal or not, is left to future work.

The data from this study will continue to be used for follow-up work. Specifically, it will be used to determine whether lurking is related in any way to the diversity of topics within a DL (i.e., breadth vs. depth of the DL), to the distribution of contributions by members of the DL (i.e., the role of stars in a DL), to the response members receive when they delurk, and to the length of messages.

Another area worth pursuing, but perhaps outside of this data set, is the investigation of high-traffic DLs and their members. For example, how do members cope with high traffic levels?

ACKNOWLEDGMENTS

Any project of this type requires a wide range of talents. We are indebted to Thawatchai Piyawat and Dick Seabrook for their consummate skills in constructing the Perl scripts and SQL database. Drafts of this paper were considerably improved through the thoughtful comments and editing of Carolyn Davidson and Heather McDonald.

REFERENCES

1. Jargon Dictionary. Lurker Definition. Available at http://www.netmeg.net/jargon.
2. Katz, J. Luring the Lurkers. Available at http://slashdot.org/features/98/12/28/1745253.shtml.
3. Kollock, P., and M. Smith. Managing the virtual commons: cooperation and conflict in computer communities. In *Proc. Computer-Mediated Commun-ication: Linguistic, Social, and Cross-Cultural Perspectives*, edited by S. Herring. Amsterdam: John Benjamins, 1996, 109-128.
4. L-Soft International. List search. Available at http://www.lsoft.com/lists/list_q.html.
5. L-Soft International. List with 1,000 subscribers or more. Available at http://www.lsoft.com/scripts/wl.exe?XS=1000.
6. L-Soft International. LISTSERV -- The mailing list management classic. Available at http://www.lsoft.com/listserv.stm.
7. Mason, B. Issues in virtual ethnography. In *Proc. Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities* (Edinburgh, 1999), 61-69.
8. McMillan, S. J. Who pays for content? Funding in interactive media. *Journal of Computer-Mediated Communication* 4, 1, (1998), Available at http://www.ascusc.org/jcmc/vol4/issue1/mcmillan.html.
9. Nonnecke, B. *Lurking in email-based discussion lists.* Ph.D. thesis: SCISM. London: South Bank University, 2000.

10. Nonnecke, B., and J. Preece. Persistence and lurkers: A pilot study. In *Proc. HICSS-33* (Maui, Hawaii, 2000), IEEE Computer Society.
11. Nonnecke, B., and J. Preece. Shedding light on lurkers in online communities. In *Proc. Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities* (Edinburgh, 1999),123-128.
12. Parks, R. M., and K. Floyd. Making friends in cyberspace. *Journal of Computer-Mediated Communi-cation* 1, 4, (1996), Available at http://www.ascusc.org/jcmc/vol1/issue4/parks. html
13. Preece, J. Empathic Communities: Balancing emotional and factual communication. *Interacting with Computers: The interdisciplinary Journal of Human-Computer Interaction* 5, 2, (1998), 32-43.
14. Preece, J., and K. Ghozati. Empathy online: A review of 100 online communities. In *Proc. Association for Information Systems, 1998 Americas Conference* (Baltimore, MD, 1998).
15. Roberts, T. L. Are newsgroups virtual communities? In *Proc. CHI'98 Conference on Human Factors in Computing Systems* (Los Angeles, CA, 1998), ACM Press, 360-367.
16. Sternstein, M. *Statistics: Barron's Educational Series*, Inc., Hauppauge, NY, 1996.
17. TECHWR-L. About the Digest Reader. Available at http://www.raycomm.com/tech whirl/digestreader.html.
18. Whittaker, S., and C. Sidner. Email overload: Exploring personal information management of email. In *Proc. CHI'96* (Vancouver, BC, 1996), ACM Press, 276-283.
19. Whittaker, S., L. Terveen, W. Hill, and L. Cherny. The dynamics of mass interaction. In *Proc. CSCW'98* (Seattle, WA, 1998), ACM Press, 257-264.
20. Worth, E. R., and T. B. Patrick. Do electronic mail discussion lists act as virtual colleagues. In *Proc. AMIA Annual Fall Symposium* (1997), 325-329.

Appendix F: HICCS 2000 paper

Appendix F contains the following refereed conference paper:

Nonnecke, B., & Preece, J. (2000). Persistence and lurkers: A pilot study. *Proc. HICSS-33 Con.*, Maui, Hawaii.

Persistence and Lurkers in Discussion Lists: A Pilot Study

Blair Nonnecke & Jennifer Preece
Information Systems Department, University of Maryland Baltimore County
1000 Hilltop Circle
Baltimore, MD 21250
{nonnecke, preece}@umbc.edu
Nonnecke: (410) 233-5771
Preece: (410) 455-6238

Abstract

In email-based discussion lists (DLs), messages resident in archives, email clients and elsewhere are persistent. One way of examining persistent messages is through the eyes of lurkers. For participants in this study, persistent conversation is an inhibitor to participation, a mechanism for engendering participation, and something to be managed. Participants viewed persistent conversation, especially when it can be retrieved through search mechanisms at a later date, as a loss of security and privacy, and an impediment to public participation. Participants followed conversations to understand the practices and language of a DL. Strategies for reading and managing email were idiosyncratic and goal driven. Some participants were concerned about maintaining access to DL information for future use. Others, more concerned about being overloaded with too much email, focused on eliminating messages. Where possible, design implications are put forward.

1. Introduction

Although lurkers are in the majority in DLs [4, 8], little research has been forthcoming on what they do and why. Apart from several academic works [5, 6] and an online column [4], lurkers in online discussion groups have received little attention. To better understand lurkers and lurking, we are developing a model of lurking that includes the following elements:

- personal forces within the individual lurker
- characteristics of the medium of email
- characteristics of email based discussion groups
- characteristics of dialogue within a DL
- tools and their influence
- activities of lurking.

Persistence as it affects lurkers falls under most of these elements, but most obviously as a characteristic of the medium of email. This paper concerns itself with understanding how persistence affects lurking and in particular how lurkers view it as a benefit to lurking, as a hindrance to public participation, and as overhead.

To put the discussion of lurkers in context, the following definitions may be helpful. The Jargon Dictionary [3] defines lurker as:

> One of the 'silent majority' in a electronic forum; one who posts occasionally or not at all but is known to read the group's postings regularly. This term is not pejorative and indeed is casually used reflexively: "Oh, I'm just lurking." When a lurker speaks up for the first time, this is called 'delurking'.

In contrast, Merriam-Webster's WWWebster Dictionary [9] provides a pejorative definition for lurk:

> **a :** to lie in wait in a place of concealment especially for an evil purpose **b :** to move furtively or inconspicuously **c :** to persist in staying

These definitions provide divergent perspectives and reflect an inadequate understanding of the lurker. The Jargon Dictionary definition evokes the image of a benevolent yet responsible Net citizen, while the traditional definition implies something much more sinister.

For the purpose of this article, a lurker in a DL is defined as anyone who for prolonged periods receives communications without publicly posting. For many this may mean never posting in some DLs. We view lurking without a pejorative or negative connotation, but as a means of participating in a DL, albeit without public posting. The goal of this research is to improve our

collective understanding of lurkers, their activities, and the tools they use. Persistence is a key element in framing our understanding of the lurker.

1.1 Persistence of email in DLs

L-Soft, one of several suppliers of email list management software, indicates there are over 55 million list members and over 150 thousand public and private lists using their software [7]. Their figures imply that DLs are widely used and cover a broad range of topics.

DLs facilitate delivery of email to a set of subscribed members using a broadcast model. Anyone who sends email to the central server effectively broadcasts the email to all members of the DL. Individuals can respond to received email via the server, which in turn broadcasts the reply to all members. There may be an intermediate step in which messages are moderated. This can introduce delays in propagation and/or the elimination of some email depending on how the moderation is handled. At the member's option, the email may be received individually or in digest form (containing a group of messages). For the most part, DLs are automatic devices for sending and receiving messages amongst members. DLs are asynchronous communication tools in which members can choose when to view their messages, if at all.

As persistent correspondence, email from DLs may be found in many places, some private, some public, and some corporate. These include the receivers' computers, of which there is typically one per member, and often times in a central archive. Email may also be intercepted or backed up and held in corporate databases. Being both persistent and dispersed means DL messages are searchable and manipulable by many.

Whether as single emails, digests, or archives, DL messages contain header information that includes sender, date, and subject. The header information also allows messages to be sorted and managed through various software, e.g., following a thread by sorting messages by the author, subject, and date headers in an email client such as Eudora. In addition, each message contains content and frequently a signature. Given the ubiquitous residences, the messages make great fodder for searching. Searching can range from a member searching her/his own locally maintained email to a researcher searching for quoted text through the use of crawler-based search engines, e.g., Excite [1].

Because copies may reside in many locations outside of the subscriber's control, access is effectively wide open. The messages can be searched for content, originator, or in many other ways. DL email may be read in isolation and the flow and intent of the messages can be distorted through the redistribution of individual messages or parts of copied message. For all intents and purposes, email from DLs may be mutated from dialogue to data and back to content, without the originator having control over the process or use. Examples include the trolling of DLs for the creation of address lists, which are then sold to spammers and legitimate businesses. The copies may also be used in the way they were intended, e.g., as an accessible resource for the group, for finding specific information, and for following conversations.

A number of factors affect how individuals manage their DL email, e.g., total number of messages, their rate of receipt, age, and size. DLs vary in the quality of messages, the number of members, content, topic, terminology, etc. All affect the ways in which the messages are dealt with by the receiver, and thus the level of persistence the messages have at a local level.

In addition to the primary artifacts (the messages, digests and archives), there are often related Web sites, sets of frequently asked questions (FAQs), and electronic forums such as chat rooms and bulletin board systems (BBSs). The principal focus of this paper is on persistence as it relates to the primary conversational artifacts.

1.2 Design issues

The persistent artifacts of DLs, the email messages, digests and archives are viewed and manipulated with a variety of tools, from UNIX mail to integrated Web browser-email-news reader tools. Archives themselves can be downloaded as text or, as is frequently the case, imbedded in a Web-based user interface (UI). As part of a larger study, the authors' current interests lie in email clients and not the Web-based UIs to archives.

Email clients have various levels of functionality and usability and have changed over the years, the most obvious differences being found in the graphical user interfaces (GUIs). There is little evidence of email clients being designed for the purpose of dealing with email from DLs, i.e., they were not designed for the volume of messages or the variety of activities associated with belonging to a DL or multiple DLs.

The differences in email clients, and the fact that these tools are being used for the more general purpose of sending and receiving email, means

that no two users will see DL emails or act upon them in an identical fashion. For example, while some users will use filters to direct messages to DL-specific mail boxes, others deal with all messages in their primary inbox. In the latter case, DL messages are intermixed with other types of email. It is at the email client and to a lesser degree the server software, where many of the design implications will be discussed.

There are design issues for other areas such as server administration. How a DL owner administers a DL has an affect on the lurking process. For example, DL owners determine whether DL content is delivered as a digest or as separate emails. Lurkers themselves have a profound affect on the quality of their experience; individuals with high levels of expertise with their email client may be better able to receive higher volumes of email or capable of belonging to a greater number of DLs.

2. Method

Our goal in this work was to develop a preliminary understanding of lurkers and lurking. As such we were looking for a method that would reveal as much as possible in as rich a way as possible. Our method borrows from Nielsen's discount usability [10] in that we were interested in gaining as much insight as possible given our resources. It also draws from Contextual Design [2] in the way the information was gathered and analyzed. This pilot study forms the basis for a more detailed examination of lurking in DLs. Results from a follow-up demographic study can be found in Nonnecke [11].

Lurkers have received very little attention, and as this was our first study of lurking, we assumed, rightly or wrongly, that lurkers would not respond to either email or Web-based surveys. We also surmised that if lurkers did respond we would get a biased response without easily understanding the nature of the bias. More importantly, we were interested in a richer understanding of lurking than could be derived from surveys. However, email and Web-based surveys have shown potential in other studies [15], and will likely be employed in our follow-up work.

Given the relatively high incidence of lurkers, we felt comfortable selecting participants at random from physical communities in which members were known to be Internet users. Ten interviewees were drawn from two locales, 5 men and 5 women, ranging in age from early 20s to early 50s. Our intention with the small sample size was to balance for age and gender, rather than examine age or gender issues. All participants were members of at least one online group, and were not pre-selected for lurking or for their level of experience with online communities. All persons asked participated in the study; 3 were well known to the researchers, and 7 were not.

Face-to-face or phone-based, open-ended interviews lasted between 45 minutes and 2 hours, and focused on the interviewee's participation in online groups. Prompting was minimal, and the interviewer did not validate whether a group or topic was worth discussing.

3. Results and Discussion

The participants described 41 communities of which there were 25 DLs, 7 BBSs, 5 newsgroups, 3 chat rooms, and 1 MOO. All participants belonged to or had belonged to communities in which they never posted, or posted rarely, e.g., once or twice, or so infrequently that they considered themselves to be lurkers. All participants had posted in at least one of their online communities.

The participants in this study were all aware of issues surrounding persistence of email without necessarily having thought about persistence as a factor affecting their actions. Many of the comments made during the interviews were related to other issues, and not persistence per se, e.g., time available, minimizing effort, privacy, entertainment value, and searchability. Recognizing that many of these issues are related to persistence, we examined our interview data from the perspective of persistence, breaking the analysis into three primary areas as it relates to DLs:

6. persistence as an aid to lurking
7. persistence as an incentive to lurk
8. persistence as overhead to lurking.

These areas were chosen for several reasons. With respect to the first area, we view lurking as a form of participation, and want to understand how persistence affects lurkers in a positive way, i.e., as an aid to lurking. In a previous paper [12] we described lurking as a means of satisfying goals. However, we recognize that lurking is not merely goal satisfaction and that other influences are at play. Our second area represents an attempt to understand how persistence acts as an incentive to lurk. The third area reflects our understanding that lurking is work, and like any other activity has overheads associated with it. If we can identify these overheads, then we may better understand the design implications.

3.1 Lurking Experience

To provide a flavor of the participants' lurking, two participants' experiences are presented. The first describes the impact of receiving messages in digest form and the second describes the general lurking process of a participant.

Cathleen's lurking using digests. Cathleen is a well known member of high standing in a professional organization. She is also a very private person. Having a health problem, she sought out and subscribed to an online DL specializing in her health problem. She read and saved all messages in digest form for several months but found it difficult to follow threaded conversations. Some digests were printed to aid reading and to save important messages. After the initial period Cathleen started skimming the digests, reading individual posts based on the subject heading. As her health improved, the DL became less important to her. Throughout her membership, she found the moderation to be aggressive and disruptive. She was eventually removed from the DL by a moderator who falsely accused her of cross-posting.

Fred's general lurking strategy. Fred is a knowledgeable, long-term user of DLs and is a technically sophisticated group member. He has belonged to a variety of DLs for reasons ranging from professional to personal interest. As a general rule, he does not post to DLs, preferring to post directly to individuals based on their public posts. He belongs to personal interest DLs to learn about the communities and for entertainment.

Fred follows threads but does not read every message in the thread. If he is very busy, he will delete messages without reading them, confident that the same issue will arise at a later date. When investigating a particular message, he uses the subject header and reads the first paragraph before continuing on. He reads to discover others' problems (e.g., technical problems with software), and says it is difficult to find this type of information in any other way, i.e., it is hard to ask about a specific problem when you don't know the problem exists.

He is also interested in learning about the community, stating that learning about the members helps him to learn about the community. He systematically described his method of coming to know members:

- information is gleaned from email address, name, signature, and URL
- understanding the members comes from what each says and how it is said
- inferences can be drawn from the choice of a false name
- knowledge about posters' habits comes from their frequency of posting and the time of day they post.

On joining a new DL, Fred reads every message to get a broad sense of the DL. He looks for cross-posts as they tell him how members view the DL in the context of related DLs and newsgroups. DL rules describing topics, moderation policy, and membership requirements, etc. tell him much about the community. Likewise, what members say and how they say it is also informative.

Each of the following three sections contains a series of lurkers' goals drawn from the participants' description of their lurking in DLs. For each goal design implications are described.

3.2 Persistence as an aid to lurking

This section describes how persistence was found to help lurkers:

- understand the DL
- satisfy personal needs
- satisfy information needs.

Goal: Understand the DL. Participants described the process of understanding the DL as a period of intense reading of most, if not all, posts. This occurred whether the posts were available as separate emails, digests, or archives. In several cases, individual posts were supplemented by searching and reading archives. During this period, which ranged from days to months, participants worked at identifying the topic or topics of the DL and determining whether this was a good fit for their needs.

Participants also worked at understanding the character of the DL. They did this to increase their understanding of the DL and to become more comfortable with the possibility of submitting messages to the DL, or in several cases side posting to individuals. Character is used very loosely here, and includes:

- terminology or special language
- posters (players and archetypes)
- rules (implicit and explicit)
- responsibilities related to being a member of the DL (implicit and explicit)
- style(s) of interaction, e.g., confrontational, humorous, etc.

- response of members to delurkers
- style and intrusiveness of moderation
- response time to messages
- volume of postings.

Participants' intentions with regard to public posting generally varied from DL to DL. For example, a participant may have joined a DL with the intention of observing and never posting from the outset. If there was a mismatch between their expertise and that expressed in the DL, then this initial period of lurking was used to confirm this before unsubscribing or remaining subscribed but completely uninvolved in the DL. If there was a possibility of posting publicly, they used this period to gauge whether their posts would have value. They also observed whether they would be able to add value, and whether the value, they could contribute, already existed in the DL, i.e., postings by others would make their contribution redundant.

New DL members are inquisitive and DL owners need to take advantage of this. The following are some of the information types that could prove valuable:

- terminology dictionary
- rules, if any
- selected highlights from the archives
- selected personal stories, e.g., in health support DLs
- description of moderation (and moderators)
- topic lists
- message rate
- number of active posters
- number of members.

While some of the above are provided in the subscription or welcome message, messages of this type were frequently unread by the participants. Key to the success of this type of information is making access obvious, timely, and ubiquitous. Many of the above list items could be kept in a DL-related Web site. A link to the Web site appended to each DL message could provide access to the site. Unfortunately, having a related Web site and linking back to it is not widely practiced.

Goal: Satisfy personal needs. When DLs were joined for personal reasons there was a correspondingly strong motivation to get as much out of the DL as possible. Entertainment was a common theme and took a variety of forms. Just as some people enjoy receiving snail mail, several participants enjoyed receiving email, indicating they liked having new email in their inbox. This gave them a sense of connection and also something to do in their free time. Others mentioned being attracted to controversy and debate, including watching flaming from the sidelines. Humour was also appreciated. Curiosity and learning were high on many peoples' list of reasons for joining and lurking in a DL.

Others joined DLs with many of the same members as their non-electronic based organizations. In their opinion, this complemented and strengthened relationships. DLs also provided a convenient way to track events and announcements. One participant, who belonged to such a DL, read all messages and deleted all but the announcements for physical meetings.

Some participants are attracted to health-support DLs as a source of empathy [14]. For at least one participant, empathy was strongly felt while lurking. DLs can also act as a mechanism for putting people in contact with one another through more private channels. For example, peers, expertise, and finding people beyond a local geographic community were described as reasons for joining a DL. Topics of specific interest to participants also drew them into joining DLs. Participants often described members of DLs as interested and focused. Relationships developed out of belonging to the DL, although no long-lasting friendships were reported as found elsewhere [13]. Several participants indicated they developed a sense of community through lurking.

For some, persistence in the form of archives provides a sense of security, allowing them to search or review a full set of messages. Having access to such an archive appears to reduce the fear of missing something.

Members of DLs have a variety of personal needs to satisfy. These are far ranging and a number of different approaches could be taken to improve and ensure they are satisfied. These include:

- providing profiles of members (to enable contact between individuals)
- suggesting related DLs and organizations, indicating attributes and differences
- providing sets of personal stories in health-support communities.

Obtaining the above types of information and keeping it current may be more of a challenge than making it available in a usable fashion. Profile information may not be necessary for all DLs, and unless there is a proven need, may require more effort that it is worth, i.e., to collect, maintain and ensure against its misuse.

At the operational level, a means of identifying specific types of messages, e.g., announcements, moderator comments, obvious flames, would aid the lurker in sorting and using the messages more effectively. A number of DLs already employ subject header prefixes for identifying message types. For this to work, members must comply with the conventions or have a moderator determine each message's type. Knowing the conventions requires an educated poster.

Goal: Satisfy information needs. Satisfying information needs was important to the participants. In some cases, information was more important than interaction. Having information in the form of archives was useful, especially if it was readily searchable. In a more passive way, the turnover of information through members' dialogue was also informative. In this way, participants were able to identify experts and if need be, seek expertise directly from these individuals.

Participants sought three types of information: factual information (e.g., job postings, and solutions to technical problems); different viewpoints arising from different levels of expertise; and access to personal experiences of others. Participants also mentioned breadth and depth of expertise as being important, as was finding "authentic" information based on an individual or group experience. Timely information was also considered quite important both in the sense of it being current, and that it meet the participants' immediate needs. Getting information from people living in the Middle East during the Gulf War was given as an example of timely information.

Professional needs, such as keeping abreast of conferences and work being done by peers and colleagues, were cited. Understanding who is doing what and where appears to be an important part of staying abreast of a professional community, particularly an academic one.

Artifacts and mechanisms for satisfying information needs must be better understood and their UIs improved. DL archives should be considered as information resources and their UIs should be designed to exploit this. Individuals within a DL act as living information sources; identifying expertise within a DL and making this identification known to members would aid information seeking. As in the previous set of design implications, message typing would be valuable for information seeking, e.g., to identify profession-related announcements.

3.3 Persistence as an incentive to lurk

How does the persistence of the DL artifacts (email, digest and archives) affect public participation in DLs? What are the issues for lurkers related to persistence that keep them lurking when they might otherwise be willing to participate?

Goal: Ensure privacy and safety. Participants were generally aware that DLs have a life of their own, and that the combination of persistence and later uncontrolled access means that there is no such thing as privacy. This inhibited their posting of personal information, and in one case, a participant's employer prohibited posting. Privacy is a concern not only at the time of the posting, but also as a long term consideration due to the persistence of DL artifacts.

Members and potential members of a DL should have a clear understanding of the implications of posting, i.e., loss of privacy. Part of that understanding lies in knowing whether the DL is publicly archived, whether there are membership criteria that have to be met in order to join the DL, and whether a list of members is readily available. At this time the majority of DLs do not provide membership lists [11].

Safety is also a concern for some lurkers. Participants who had concerns about safety expressed it at two different levels. The first relates to a fear of violence, i.e., that someone or some agency can use posted information (or mere membership in a DL) to find someone or something about someone. The second relates to the fact that if you don't post you can't offend, and therefore will not become a target of flaming. While the safety issue is different from privacy, the design implications are similar.

One option for ensuring privacy and safety is the use of anonymous email hosting services such as hotmail.com. These services provide mechanisms for anonymously posting and receiving messages. There is a conundrum; participants were interested in maintaining their own privacy yet wanted to know more about other members. For example, a poster's address and signature were mentioned as a means of understanding the poster, and one participant wanted to find DL members of a similar age and gender.

Goal: Reduce noise and exposure. Most participants realized that DLs and other online forums are regularly pilfered for email addresses, which are then sold or used directly to spam. Not

one participant said they look forward to receiving spams. Spammers can obtain messages directly from the messages themselves or by querying the DL server for a list of members.

As a first level of defense members' addresses should be made difficult to access. Owners of DLs can easily restrict access to the DL membership list. Similarly, DL server software can be set up to prevent the distribution of email from non-members. Some DL members take their protection one step further and provide incorrect return addresses on their email. While this may foil spamming, it makes legitimate communication difficult, e.g., to get the correct address takes more effort when side posting.

DLs allow emotional detachment as the audience and thus the lurkers are for the most part not identifiable. As one participant expressed it, when you lurk, you can have curiosity without exposure. In contrast, several participants indicated that it is much more difficult to lurk in chat rooms than DLs as chat rooms are synchronous environments where participants are normally visible and thus approachable. For some participants, the practice of lurking makes leaving a DL easier in that there is less of a commitment to a DL if you don't post. For some individuals, their notoriety makes posting problematic, e.g., few government officials post to public DLs.

Some DLs discourage lurking, at least at the outset, suggesting in their introductory message that newcomers should provide a description of themselves and post it to the DL. Other DLs specifically state that posting is not required. In either case being aware of the rules of the DL is an important part of participation. Few of the participants in this study indicated that they read the rules or guidelines.

3.4 Persistence as overhead to lurking

As used in this paper, overhead is defined as the set of actions and time required by the lurker to deal with DL email. For all participants, DL email was received along with other email through a single preferred email client. These email clients varied in type and configuration for each participant. As such, each participant received DL email under very different conditions. To add to the variety in overhead, their skills with the email clients ranged from naïve to expert, and the tasks they performed ranged from simple to complex.

Participants had other priorities in their lives; DL reading/following was frequently not the most important task of the day and certainly not the one in which they wished to spend most of their time, or even a good portion of their time. In the context of their lives, lurking in a DL is one of many activities filling their day. The following is a synopsis of participants' goals, their overheads and the resulting design implications.

Goal: Maximize return. In general the participants were interested in getting the most out of the time they had for lurking. Even if they lurked to entertain themselves, they still wished to do this as efficiently as possible. This typically meant spending less rather than more time with the DL(s).

They used a number of methods to do this. If they belonged to more than one DL, they limited themselves to the number of DLs they could handle. It was clear that too many DLs meant that the value of one or more of the DLs would be reduced.

While many of the DLs described by participants had 20-30 messages/day, participants were generally happier with fewer message. Factors affecting the amount of time required to lurk on a DL include the quality and size of the messages, the motivation in belonging to the DL, the volume and type of email received from other sources, and the time available. In our examination of a number of introductory messages and DL related Web sites, none mentioned how many messages a subscriber might expect.

The asynchronous and persistent nature of DLs means that lurkers can go back through older messages at any time and either search for particular information or browse the messages.

Goal: Keep inbox manageable. Manageable meant different things to different participants, but was often related to comfort. For several participants comfort came from keeping their inbox small, i.e., able to see all retained messages at once. The process of picking through the messages was an important part of their management process. Understanding how inboxes are used is critical to developing design solutions.

The use of filters to sort messages into secondary mail boxes was not commonly used among participants. A number of reasons were stated: not trusting the effectiveness of the filters, potential burying of important email, and no knowledge of filtering tools or the process of creating effective filters. Filtering mechanisms should be examined with an aim to making them verifiable, trustworthy, and simpler to learn and use.

Goal: Identify DL email amongst other email. Differentiating one DL's messages from another, and those in turn from non-DL email was an effort for participants. Recognizing this as a problem, some DLs use an identifying prefix in the subject header to indicate that a message is from a particular DL, e.g., the MORE cycle DL prefixes all subject headers with "more:".

Identification of DL messages is an important mechanism for scanning and processing email in the inbox and elsewhere. The current ad hoc approach of using prefixes may be good enough, but could be improved upon. A related issue, although not raised by the participants is the use of prefixes to identify different types of messages, e.g., "Q:" for question. The use of prefixes helps identify a message's origin and intent, but may also make the subject heading more difficult to read.

Existing header information is sufficiently descriptive for use in separating messages from different DLs and non-DL email. As mentioned in the previous goal, filtering tools remained largely unused by the participants. Whittaker and Sidner [17] found the inbox to be an important repository for messages. Their findings suggest that the low use of filters may not reside solely in the act of filtering, but on other factors, such as the fear of losing track of important information.

Goal: Follow threads. A thread is a conversation of multiple messages linked via a repeatedly used subject header. Participants were able to follow threads in newsgroups and BBSs because these systems were designed with threaded conversation in mind. Participants used threading to either follow a particular discussion or determine whether a line of discussion was worth reading. This particular facility is poorly implemented or non-existent in most email clients. In addition, threading in email clients is different from that in newsgroups or BBSs. Even when messages can be sorted by subject header in an email client, the results are presented as a list of messages related by subject header. In both BBS and newsgroups, messages are related in a tree like manner, with the relationships between individual messages being apparent to the user. For this reason, email-based threading might better be called clumping.

For threading to be of value in email clients, threading must be effectively represented in the UI, e.g., threading based on subject header and date, and keeping the most active threads in the most visible position in order for the thread activity to remain observable to the user. GUI-based email clients can show threading based on the subject headers, but the results are frequently cumbersome and confusing. Alternative solutions need to be examined.

Additional problems occur when receiving DLs as digests. While this reduces message clutter in the inbox, it eliminates thread visibility. Current email clients are unable to show threading in digests although specialized digest readers such a Digester [16] show promise in this area.

Goal: To read or not to read. Determining what to read is an important activity for any lurker. Deciding whether a message was worth reading was idiosyncratic and for a given participant often differed between DLs. The following criteria were described:

- read all if participant is new to the DL
- read if the subject heading shows potential value
- read if the author is known
- read all messages in a thread if the middle message of a thread is interesting
- read messages if thread is long (i.e., quality of messages and thread is somehow related to the length of the thread)
- read messages with confusing subjects
- read or not read an obvious flame.

Several participants deleted all or most messages (read or not read) as a matter of course whereas others kept messages, either by leaving them in the inbox and relying on the read flags to indicate their status, or by manually placing them in secondary folders. The delete process was most common among users of text-based email clients.

A rich set of cues were used in deciding whether to read a message. The fact that messages are persistent and asynchronous, means that a message does not have to be read at the time of receipt. It also means that the decision as to whether a particular message is read will often be based on other messages, e.g., other messages in the thread or the quantity of messages in the inbox.

3.5 Summary

In the three previous sections, the goal of lurkers and the corresponding effects that persistence has on those goals was outlined. In this section, the design implications are discussed based on where change would be beneficial. The following five areas are summarized below:

- email client
- server software and administration
- alternative access mechanisms

- support information
- member

There are two leading ways in which all email clients can be improved: by showing threading, and improving filtering. Threading provides lurkers with the ability to judge whether messages are valuable, and how to deal with them. It also allows the user to follow conversations more easily. At this time, threading is poorly implemented on most email clients. Filtering has the capability of separating and thus organizing multiple DLs into separate areas and thus reducing clutter in the inbox. Filtering is readily available on most GUI based email clients, but is not frequently used.

At the server level, several improvements could make life easier for the lurker. However, some improvements negatively affect other areas. For instance, digests are intended to reduce inbox clutter, especially with high volume lists. However, thread following is compromised when messages are delivered in digest form. There is some evidence [11] that DLs that are set to send out digests to new subscribers also have higher levels of lurking. Whether this is a result of digests being less well read, more difficult to respond to, or harder to follow threads is not known.

Many DLs add a prefix to the subject header as a means of identifying messages. These prefixes may make DL identification easier, but likely obscure the actual subject header. Whether knowing that a message comes from a particular DL is more important than the subject is unclear. It will likely depend on many factors, including the volume of messages in the inbox, the rate of receipt, and purpose of belonging to the DL.

Most DL administrators prevent access to membership lists. More often than not they also prevent messages being broadcast by non-members. There is however, very little they can do to prevent the pilfering of addresses from archives. Some members have taken up the anti-spam challenge by supplying incorrect return addresses.

At the level of supporting the lurker with information related to the DL, providing an accessible, current, and usable set of information is important. Creating links to it in all outgoing messages would provide access. Within the Web site or wherever it may reside, access to an archive is an important information resource for many lurkers, particularly if they are trying to understand the nature of the DL, or looking for specific information. A usable interface should allow lurkers to browse, follow threads and search for information.

An archive can also provide the functionality for posting messages. This can either reduce or eliminate the reliance on the email client. In doing so, many of the problems described so far could be reduced, e.g., threading is usually apparent in Web-based interfaces, anonymity is frequently built into the system ensuring safety and privacy, and an archive and supporting information can be integrated into the environment.

On the down side, the user may have a different identity and potentially a different password for each DL they belong to. UIs while similar in intent between DLs will be different and will require familiarization with each. In contrast, email-based DLs utilize a single familiar UI for a given user. In addition, the email client can receive email automatically, whereas, Web-based DLs rely on the user to seek them out.

For the security conscious, the use of Web-based DLs may offer a preferred solution. However, it is unclear whether security is more important than the convenient and consistent albeit somewhat underused and noisy UI of the email client.

Email clients may be facing functionality bloat already, adding additional functionality for lurking may not be the best approach. Improved lurking may come through improving the skills of the lurker. While the current email clients may not have been specifically designed with lurking and DLs in mind, many of their facilities go unused by the lurker, e.g., use of secondary mail boxes and filtering. This is in part due to users not being familiar with the functionality of the email client, but also stems from the way in which they view and use the inbox as a central repository. As is the case with other software, DL members use the tool to the extent that fulfills their immediate needs. Its likely that the level of participation (posting or not) in one or more DLs is a function of their skills in using the email client. Other factors will also be at work, such as volume of email, personal strategies, motivation, time available, etc. By improving our understanding of strategies and the context in which the strategies develop, we should be able to come up with a better model of the lurker, and improvements in their tools.

4. Conclusion

As a means for asynchronous group communication, DLs have gained wide acceptance. This is in part due to their use of the most ubiquitous of Internet tools, the email client. We have illustrated a number of issues and design implications related to persistence.

As aid to their lurking, participants described how they followed online conversations to understand the practices and language of a DL. They talked about observing the coming out of other lurkers, and measuring the group by its treatment of new members. Several participants developed a sense of community in the process of following conversational threads. As an incentive to lurking, security and privacy are very important issues. Participants viewed persistent conversation, especially when it can be retrieved through search mechanisms at a later date, as an impediment to public participation. As overhead to lurking, participants described the process of using and maintaining their DL email. Mechanisms for reading and managing email were idiosyncratic and goal driven. Some participants were concerned about maintaining access to found information for future use. Others, more concerned about being overloaded with too much email, focused on eliminating messages while getting the most out of what they could make use of in the moment.

It is not surprising that some of the lurkers' goals lead to contradictory design implications. On one hand, privacy is a very important issue, and on the other, participants desired more information about the DL and its members. Email clients are relatively simple, well understood tools which in part accounts for the popularity of DLs. Improving in their facility as DL front ends may increase their complexity and thus compromise their broad appeal. Full featured Web based UIs to DLs hold promise for eliminating many of the problems associated with the email-based UI. Whether lurkers are willing to switch to an alternative UI is unknown.

The method of using a small group of participants, interviewing them with regard to the membership and practices within online groups was an effective technique for exposing a wide variety of issues related to lurking. We now have a good base for carrying out in-depth interviews and surveys in order to understand the relevance of the findings and the usefulness of the design implications.

The design implications coming from this work should be of interest to developers of email clients. This work will be followed by evaluations of several email clients and at least one digest reader, with an eye to examining how the tools have shaped usage, and how current usage can improve the design of the tools. While this work has focused on lurking, the issues raised are applicable to all DL members regardless of whether they post or not. There will be additional specific issues and design implications for public posting in DLs.

5. Acknowledgments

This work would not have been possible without our interviewees; we thank them for their time and insight. Carolyn Davidson and Heather McDonald provided thought provoking and valuable reviews of early drafts.

6. References

[1] Excite.com. What We Do. Available at http://corp.excite.com/Company/what.html.

[2] Holtzblat, K., and H. Beyer. Contextual Design: using customer work models to drive system design. *Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems*, 1995), 391-392.

[3] Jargon-Dictionary. Lurker Definition. Available at http://www.netmeg.net/jargon.

[4] Katz, J. Luring the Lurkers. Available at http://slashdot.org/features/98/12/28/1745253.shtml.

[5] Kollock, P., and M. Smith. Introduction: Communities in Cyberspace. In: *Communities in Cyberspace*, edited by P. Kollock and M. Smith. London: Routledge, 1999.

[6] Kollock, P., and M. Smith. Managing the virtual commons: cooperation and conflict in computer communities. In: *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*, edited by S. Herring. Amsterdam: John Benjamins, 1996, 109-128.

[7] L-Soft International. Record sites. Available at http:// www.lsoft.com/listserv.stm.

[8] Mason, B. Issues in virtual ethnography. *Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities* (Edinburgh, 1999), 61-69.

[9] Merriam-Webster. Lurk. Available at http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=lurk.

[10] Nielsen, J. *Usability Engineering:* Academic Press, Boston, 1993.

[11] Nonnecke, B. Lurking in email-based discussion lists. In: *SCISM*. London: South Bank University, 2000.

[12] Nonnecke, B., and J. Preece. Shedding light on lurkers in online communities. *Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities* (Edinburgh, 1999), 123-128.

[13] Parks, R. M., and K. Floyd. Making friends in cyberspace. *Journal of Computer-Mediated Communication 1*, 4, (1996).

[14] Preece, J. Empathic Communities: Balancing emotional and factual communication. *Interacting with Computers: The interdisciplinary Journal of Human-Computer Interaction 5*, 2, (1998), 32-43.

[15] Smith, C. B. Casting the net: Surveying an Internet population. *Journal of Computer-Mediated Communication 3*, 3, (1997).

[16] TECHWR-L. About the Digest Reader. Available at http://www.raycomm.com/techwhirl/digestreader.html.

[17] Whittaker, S., and C. Sidner. Email overload: Exploring personal information management of email. *CHI'96* (Vancouver, BC, 1996), ACM Press, 276-283.

Appendix G: Ethnographic Studies in Real and Virtual Environments paper

Appendix G contains the following refereed conference paper:

Nonnecke, B., & Preece, J. (1999). Shedding light on lurkers in online communities. *Proc. Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities Con.*, Edinburgh.

Shedding Light on Lurkers in Online Communities

Blair Nonnecke & Jennifer Preece
IFSM, University of Maryland Baltimore County
1000 Hilltop Circle
Baltimore, MD 21229
{nonnecke, preece} @umbc.edu

Abstract

Lurkers are reported to make up a sizable proportion of many online communities, yet little is known about their reasons for lurking, who they are, and how they lurk. In this study, interviews with online community members provided a formative understanding of these and other issues. We discovered that lurking is a systematic and idiosyncratic process, with well-developed rationales and strategies. All interviewees lurked, but not all the time, and several developed a sense of community through their lurking.

Introduction

Recent research in electronic discussion groups has focused on a number of areas, including the nature of online communities (Wellman, 1997), the development of friendship (Park & Floyd, 1996) , the role of empathy in group discussions (Preece, 1998), and the differences between men and women (Roberts, 1998). Additional work has been done on specific kinds of online communities, e.g., therapy (King, 1994), education (Hiltz, 1993), business (Sproull, 1986), and health support (Preece & Ghozati, 1998). In most of these studies, the primary source of information was participants who actively conversed in the discussion groups or other online forums, and who were therefore readily observable. While our knowledge is growing, it is nevertheless a selective knowledge based on observations of those who post.

In contrast, little has been published on the so-called lurkers, those who do not participate in public dialogue. Given that estimates of lurkers-to-posters ratios run as high as 100:1 (Carroll & Rosson, 1996) — possibly higher in some electronic discussion groups — a knowledge of lurkers would be a valuable addition to our understanding of "participation" and the design of electronic discussion groups. Our interest in this area originated in our research into online health support groups, where the number of lurkers tends to be very high. These unseen, but needful users deserve special attention from designers interested in supporting these groups.

To put our discussion of lurkers in context, the following definitions may be helpful. The Jargon Dictionary (http://www.netmeg.net/jargon) defines lurker as:

> One of the 'silent majority' in a electronic forum; one who posts occasionally or not at all but is known to read the group's postings regularly. This term is not pejorative and indeed is casually used reflexively: "Oh, I'm just lurking." When a lurker speaks up for the first time, this is called 'delurking'.

In contrast, Merriam-Webster OnLine (http://www.m-w.com) provides a strongly pejorative definition:

> Lurker … mean[s] to behave so as to escape attention. LURK implies a lying in wait in a place of concealment and often suggests an evil intent, e.g., suspicious men *lurking* in alleyways.

These definitions provide contrasting perspectives, and reflect our inadequate understanding of the lurker. In the online context, lurker evokes the image of a benevolent yet responsible net citizen, while the traditional definition implies something much more sinister. Evidence for the former definition is anecdotal, and without appreciating the nature of online lurking, the latter definition may be inappropriate. An improved definition, based on empirical evidence, will have practical value in defining how we view lurkers, and how electronic discussion groups can better serve the lurker.

We have identified four major areas of investigation: lurker demographics (who and where), lurker strategies and tactics (how), the effect of context (community variables), and lurker activities (why and what).

Demographics: If the estimated 100:1 ratio of lurkers to posters is accurate – and many researchers believe it may be higher – the sheer volume of lurkers makes them an important group. Unfortunately, we have neither a good count of how many lurkers are out there, nor clear reasons why lurking levels may vary from community to community. To clarify this, we are concurrently examining lurking levels based on community type (Nonnecke & Preece, 1999). Understanding lurkers and lurking offers us the opportunity to develop measuring techniques and metrics that include them, rather than treat them simply as the silent-insignificant majority. Metrics will be invaluable in measuring the success of a community, whether that be social, economic or whatever. This has already become a financially important topic as more and more commerce comes to the Net. (Hagel & Armstrong, 1997)

Strategies and Tactics: Viewing lurking as a valid form of participation in online communities suggests that we should understand the usability of the community, i.e., how communities function on lurkers' behalf. If we understand the strategies and tactics of the lurker, e.g., browsing, reading practices, side posting (if any), and mechanisms for archiving or saving information, then we have the opportunity to positively influence the tool makers and designers of these communities. We are in the process of examining strategies and tactics in detail.

Context: The communities in which lurking occurs affect participation. Factors include topic, software, size of community, number of posts, moderation (Collins & Berge, 1997), and joining policy. Coming to understand at least some of these, if not exactly knowing their generalized effect, will help us perceive lurking as a situated action, and not just an isolated phenomenon.

Activities: Do lurkers, as the Jargon Dictionary suggests, regularly read messages? Why don't they post, and what are the benefits of participating without posting? Are they a homogeneous group, i.e., do they use the same strategies and techniques, or is there a range of lurking activities and rationales? How do they view themselves: as lurkers in the evil sense, as members of a community, or somewhere in between? The research described in the remainder of this paper focuses on lurker activities.

While data from our other research efforts will provide us with a good understanding of lurker demographics, it will tell us little about lurker activities. Rather than view lurking in the pejorative, we have taken a neutral position, and view it as a form of participation. This raises the question of exactly what is lurking. While it could be assumed that someone is lurking any time they are not posting, we have temporarily defined lurking as follows:

> 'prolonged periods of receiving communications without posting. For many people this may mean never posting in some communities'.

Indeed, in discussing posting habits with lurkers, we discovered it was common for people to post in some communities and never in others. Our initial goal was to gain a formative understanding of lurking, to identify some of the factors that influence lurkers, to identify potentially important variables and to develop

tentative hypotheses to drive further studies. To understand the life of lurkers, we looked to members of online communities to describe their own participation.

Methodology

Lurkers have received very little attention, and as this was our first study of lurking, we assumed, rightly or wrongly, that lurkers would not respond to either email or Web-based surveys. We also surmised that if lurkers did respond we would get a biased response without easily understanding the nature of the bias. More importantly, we were interested in a richer understanding of lurking than could be derived from surveys. However, email and Web-based surveys have shown potential in other studies (Smith, 1997), and will likely be employed in our follow-up work.

Given the relatively high incidence of lurkers, we felt comfortable selecting participants at random from physical communities in which members were known to be Internet users. Ten interviewees were drawn from two universities, 5 men and 5 women, ranging in ages from early 20s to early 50s. Due to the small sample size, our intention was to balance for age and gender, rather than examine age or gender issues. All participants were members of at least one online group, and were not pre-selected for lurking or for their level of experience with online communities. All persons asked participated in the study; 3 were well known to the researchers, and 7 were not.

Face-to-face or phone-based, open-ended interviews lasted between 45 minutes and 2 hours, and focused on the interviewee's participation in online groups (see the following Interview Outline). Prompting was minimal, and the interviewer did not validate whether a group or topic was worth discussing.

Interview Outline:

~~1.~~7. Introduction
~~2.~~8. Name and describe online groups to which they belong(ed)
~~3.~~9. For each group, determine the following:
- reasons for joining
- activities and action they took (in posting or non-posting)
- reasons for not-posting (if this occurred)
- reasons for posting (if this occurred)
- reasons for side posting in non-public spaces (if this occurred)
- reasons for leaving the group (if this occurred)

~~4.~~10. Ask for comments on Jargon Dictionary definition
~~5.~~11. Ask for any additional comments
~~6.~~12. Thank them for participating

Results & Discussion

Our ten interviewees described 41 communities, including 25 discussion lists, 7 BBSs, 5 newsgroups, 3 chat rooms, and 1 MOO. Only one community was common to two interviewees. Two interviewees described only one community, while the maximum number of communities mentioned by an individual was 6.

All interviewees belonged to or had belonged to communities in which they never posted, or posted rarely (i.e., once or twice), or posted so infrequently that they considered themselves to be lurkers. All interviewees had posted in at least one of their online communities.

Despite the small sample size, it is clear that lurking is not a simple single behaviour, but a complex set of behaviours, rationales, and activities situated in a rich space of possibilities. Using Affinity Diagrams (Holtzblatt & Beyer, 1995), two sets of related factors emerged from the interviews: factors that lead

people to join a community, and those that affect their participation. We use these observed factors as a starting point for our discussion.

i) Interviewees described the following major reasons for joining:
- personality (e.g., curiosity, boredom, need for interaction)
- social (e.g., parallel with physical community, search for like-minded people, desire to broaden beyond local geographic community)
- professional (e.g., relationship to job)
- information (e.g., access to experts, timely information, exposure to breadth and depth of ideas)
- pleasure (e.g., entertainment, controversy and debate, humour)

ii) The following reasons were provided for not posting:
- understanding the community (e.g., audience unknown, comfort level, topic area, individuals)
- personal factors (e.g., culture of origin, motivation, time)
- no personal or practical need (e.g., able to gather information without posting, just reading, no reason to respond)
- no community requirement (e.g., no expectation or requirement by community)
- structure of community (e.g., posting not possible, part of community is non-posting: FAQ, moderation)
- information seeking (e.g., more interested in information than interaction, reading with a specific goal in mind)
- privacy (e.g., sensitivity of employer, fear of archiving, fear of spamming, i.e., junk mail)
- safety (e.g., can't offend if don't post, curiosity without exposure)
- involvement (e.g., maintain emotional detachment, makes leaving easier, shy)
- community responsiveness (e.g., delay between posting and response, non-response to posts)
- value of posting (e.g., no response required, nothing to offer, unable to add value)
- interaction mechanisms (e.g., volume of posting, user interface, anonymity)
- efficiency (e.g., not posting takes less time, others will respond, value received without cost of posting)

A simple goal satisfaction model can be used to interpret and integrate the above sets: interviewees attempted to satisfy their goals in joining a community, and these goals were variable and context dependent (e.g., style of community, need for anonymity, and many other factors affect decision to join, and remain in the community). How they satisfied their goals was also context dependent; however, in many instances it was possible to satisfy goals without posting. It became clear that this was not a simple process of reading every posting, but a complex, idiosyncratic process influenced by the individual's goals, experience, and the specific community.

In one example, an interviewee belonged to a broad range of discussion lists, having joined them for both personal and business reasons. While the motivations for joining each list was different (e.g., want to know vs. need to know), participation in the lists was, for the most part, limited to lurking. Lurking was comfortable and enabled him to attain his goals given the nature of the lists, each list having high volumes of quality postings representing both depth and breadth of knowledge. In neither group was the interviewee motivated to post for information, but took a more general wait-and-see approach.

Each of the interviewees was able to describe a method for dealing with postings. Nobody read every posting, and depending on the experience within the community interviewees might not read any of the postings. Subject headings were used as a primary means of determining what to read, and the poster's name was used, if at all, as a secondary guide. Several users deleted or ignored whole threads based on a heading, well aware that heading information was often a poor indicator of content. One interviewee said this was a reasonable strategy because information tended to get recycled over a period of time.

It would be convenient (for designers, community builders, etc.) if we could cite specific reasons for non-posting as more important than others, but that does not appear to be the case with this group of interviewees. It is clear that non-posting activities are carried out methodically and that individuals are capable of explaining not only their methods, but also the strategies they employ. Several very interesting aspects of non-posting arose, but it is too early to tell, if at all, how these fit non-posters in general:

- Users have a need to understand the community, and actively work at doing so (whether they post or not).
- Time is an important determinant of how people participate.
- Privacy and safety are important issues.
- The mechanics (user-interface, administration, etc.) of a community have an impact on participation.
- A strong sense of community can be developed without posting.

Knowing the community and the individuals in the community was so important to the interviewees that many were able to describe their tactics in detail. These included looking at previous posts by an author (using archives or other means), examining email addresses for personal or corporate information, following threads to understand the nature of the discussion and participants, and using signatures and related Web sites to find out more about posters.

The fact that a strong sense of community can be developed without posting is perhaps the most interesting in that it goes against the preconceived notion that you must be an active poster to be part of a community. A number of interviewees mentioned experiencing a sense of community while lurking.

For one interviewee, a sense of community was extremely strong. This came about through a number of avenues: the interviewee's need to find community within a self help group, the stories related within the community's Web space, private postings and responses to members of the community, and the character and nature of the dialogue within the community, which engendered a sense of trust and care. The fit between the interviewee and the community was good, and in this case the outcome was a very strong sense of community, a sense that was developed without posting. Even though this interviewee has not actively lurked in the community for over a year, there is still a sense of belonging to this community.

Interviewees were asked to describe both current and past communities. As a result, we were able to get a glimpse into why they left communities. Many indicated that a lack of time was an important element in their leaving a community. However, communities cited as largely information interchange communities (e.g., software application help groups) were frequently left because they no longer supplied information in sufficient quantity or quality. This was largely a result of communities repeating topics, and the interviewees becoming more expert in their knowledge.

We were interested in understanding how interviewees viewed lurking in general and their own lurking behaviour in particular. An initial abstract for this paper was distributed to some of the interviewees. One interviewee responded with the following comment:

> Maybe it's a sign of my own mild discomfort around being a lurker, but I found it reassuring to recognize myself and my behaviour within the continuum you describe, and to see lurking treated seriously, with both acceptance and respect. As a lurker, I'm used to observing from the sidelines and participating vicariously, and it's strangely gratifying to read an article that speaks directly to that experience. It's almost like suddenly feeling part of an (until-now) invisible community of lurkers.

This interviewee was not alone in feeling there is a stigma associated with lurking, although the degree of stigmatization varied from individual to individual. Giving lurkers recognition as valid participants (beyond

the current Jargon Dictionary) will benefit both lurkers and the community as a whole. Simple math indicates that if lurkers delurked, communities in their present form would become a chaotic message ground, perhaps mimicking many newsgroups in their level of disorder.

Conclusion

Our initial assumption that lurkers could be found by polling a physical community proved to be a good one, with all interviewees being experienced lurkers, but not necessarily all of the time, nor in all communities. Further work on the demographics should provide us with a better understanding of just how wide spread lurking is on a community-by-community basis.

While we found no anecdotal evidence of gender playing a role in lurking behaviour, this does not preclude the possibility. Our ongoing work in examining lurking demographics should provide us with at least a starting point.

With the small numbers interviewed, it is premature to create a set of guidelines for tool and community builders, but we do know that lurking is a highly active, methodical, and goal-driven process. At our current level of investigation, it appears to be idiosyncratic, however, with further studies we may be able to understand it well enough to inform tool and community builders with either guidelines or a set of evaluation heuristics aimed at creating better communities for lurking.

Our refinement of the lurker definition appears to be sound: prolonged periods of receiving communications without posting, which for many may mean never posting in some communities. We will need to incorporate into this definition some elements about the sense of community, but this will have to wait until we have a better understanding of how this sense of community is achieved, and how frequently. No longer can we assume a lack of a community in cases where the majority of members do not post.

So far, we've viewed lurking as a means of satisfying a set of user-defined goals related to joining an online community. We also need to understand the circumstances under which members of online communities post publicly and privately within the community. And we need to understand why such participation ends. In a sense, we are looking at understanding the life cycle of participation in an online community, with an emphasis on lurking as being a natural and valuable part of that life cycle. Understanding why someone delurks, and its value to the community, is equally interesting as why they lurk. Understanding why someone leaves a community will be as interesting and useful as understanding why they became members in the first place.

Lurkers and lurking will continue to be an important area of study as more and more communities go online. Our next step will be to develop better tools for the lurker, thereby creating better communities for all participants. To achieve this, we need to broaden our definition of participation and take up the challenge of studying participation in all its forms through combined ethnographic and large scale sampling approaches.

Future Work

This formative study is continuing to spawn work within our group. A demographic study of online discussion groups is now being completed and will provide us with a knowledge of lurking levels. Further ethnographic studies are proposed to discern how current tools, specifically email clients, affect lurking. These studies will tell us more about the strategies and tactics employed by users, with an eye towards understanding how current tools have shaped use, and how to improve them.

The role of moderation is currently being examined with a focus on providing a better suite of tools for moderators. In particular, we are developing a technique for automatically identifying inappropriate messages (e.g., flames, spam, abusive or obscene language). This technique will eventually form the basis

for developing a tool for moderators of large communities and discussion groups. It will flag messages that are potentially inappropriate, so that the moderator need read only those messages and decide how to handle them, instead of reading every message. It is likely that the idea can be adapted and made more generic, so that moderators who, for example, wish to keep a particular discussion on track by discouraging messages that wander from the theme, can use the tool to identify those messages too.

Acknowledgments

This work would not have been possible without our interviewees; we thank them for their time and insight. Carolyn Davidson, Dave Hendry, and Richard Jacques provided thought provoking and invaluable reviews of early drafts.

References

Carroll, J. M., & Rosson, M. B. (1996). Developing the Blacksburg electronic village. *Communications of the ACM, 39*(12), 69-74.

Collins, M. P., & Berge, Z. L. (1997). Moderating online electronic discussion groups. Paper presented at the *American Educational Research Association Annual Meeting*, Chicago, IL.

Hagel, J., & Armstrong, A. G. (1997). *Net gain: expanding markets through virtual communities*. Boston: Harvard Business School Press.

Hiltz, R. (1993). *The virtual classroom: learning without limits via computer networks*. Norwood, N.J.: Ablex Publishing.

Holtzblatt, K., & Beyer, H. (1995). Contextual Design: using customer work models to drive system design. Tutorial presented at the *Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems*.

King, S. A. (1994). Analysis of electronic support groups for recovering addicts. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century*, *2*(3), 47-56.

Nonnecke, B., & Preece, J. (1999). Lurker demographics: counting the silent. *TBA*.

Park, M., & Floyd, K. (1996). Making friends in cyberspace. *Journal of Computer-Mediated Communication, 1*(4).

Preece, J. (1998). Empathic Communities: Balancing emotional and factual communication. *(In Press) Interacting with Computers: The interdisciplinary Journal of Human-Computer Interaction, 5*(2), 32-43.

Preece, J., & Ghozati, K. (1998). Empathy online: a review of 100 online communities. Paper presented at the *Association for Information Systems, 1998 Americas Conference*.

Roberts, T. L. (1998). Are newsgroups virtual communities? talking on the net. Paper presented at the *Proceedings of ACM CHI'98 Conference on Human Factors in Computing Systems*.

Smith, C. B. (1997). Casting the net: surveying an internet population. *Journal of Computer-Mediated Communication, 3*(3).

Sproull, L., and Kiesler, S. (1986). Reducing social context cues: electronic mail in organizational communication. *Management Science, 32*(11), 1492-1512.

Wellman, B. (1997). An electronic group is virtually a social network. In S. Kiesler (Ed.), *Culture of the Internet* (pp. 179-208). Mahwah, NJ: Lawrence Erlbaum Associates.