Use of the Internet as a data collection tool: a methodological investigation of sampling and mode effect

Alison Ruth Evans

Thesis submitted for the degree of
Doctor of Philosophy

City University
Department of Sociology and
School of Nursing and Midwifery

January 2006

Volume 1

Table of contents

Volume 1

Volume 2

PART III –
MODE EFFECTS IN STUDIES OF SEXUAL BEHAVIOUR

List of tables

Volume 1

Volume 2

List of figures

Volume 1

Volume 2

Acknowledgments

I would like to express my sincere thanks to my supervisors, Dick Wiggins and Jonathan Elford, for their constant support, guidance and commitment to this PhD project. My thanks also go to the members of my advisory board for their many and various contributions (Graham Bolding, Mark Davies, Graham Hart, Roger Jowell and Lorraine Sherr); to Cath Mercer for her invaluable help with the Natsal 2000 dataset; to all the participants from the Internet and HIV study, with special thanks to those who took part in my interviews; and to my family, friends and colleagues for all their words of encouragement. Thanks also to the Economic and Social Research Council who provided the funding to make this study possible.

Declaration

I grant powers to the University Librarian to allow this thesis to be copied in whole or in part without further reference to me. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Alison Ruth Evans

Abstract

This methodological investigation set out to examine the use of the Internet as a data collection tool: its potential for sampling gay and bisexual men and gathering data on sexual behaviour.

The research took place in the context of the Internet and HIV study, an examination of high risk sexual behaviour among gay and bisexual men seeking sex on the Internet. The Internet and HIV study collected quantitative data using self-administered questionnaires completed online or by pen-and-paper, and qualitative data using one-to-one interviews conducted face-to-face or online. Its design provides an invaluable opportunity for investigating Internet-based research methods.

Methodological issues were explored through collection of primary data from respondents who had participated in the Internet and HIV study and secondary analysis of data collected for the study.

The first area of exploration was the use of web surveys for sampling gay and bisexual men. The results indicated that participation in the Internet and HIV web survey was driven by an altruistic desire to contribute to research into sexual health promotion and HIV prevention. Drop out was most likely to occur in the early stages of the survey and minor differences in sub-group drop out were identified. Men who participated in the web survey were broadly similar to the MSM drawn from a national probability sample, but were more likely to participate in high risk sexual activity. Differences were more pronounced when the samples were restricted to London men.

The second area of exploration was the effect of the Internet mode on data collection. The web survey reduced item nonresponse in comparison to the pen-and-paper survey but did not improve response to sensitive questions over and above this general mode effect. The results indicated that online synchronous interviews may be suited to a more structured format, given the difficulty of expressing complex ideas through typing in real time.

It was concluded that the Internet has excellent potential for researching sexual behaviour among gay and bisexual men. It provides effective access to this population and more complete survey data than the traditional alternative. The application of online interviewing requires further exploration.

Abbreviations

ACASI	audio computer assisted self interview
AIDS	auto immune deficiency syndrome
BIDR	Balanced Inventory of Desirable Responding
CAPI	computer assisted personal interview
CASI	computer assisted self interview
CATI	computer assisted telephone interview
DBM	disk-by-mail
ESRC	Economic and Social Research Council
HIV	human immunodeficiency virus
HTML	hyper text markup language
IRC	Internet relay chat
KMO	Kaiser-Meyer-Olkin measure of sampling adequacy
LGB	lesbian, gay, bisexual
LGBT	lesbian, gay, bisexual, transgender
MEG	moderated email groups
MSM	men who have sex with men
MOO	MUD object oriented system
MUD	multi-user dungeon
Natsal	National Survey of Sexual Attitudes and Lifestyles
NHS	National Health Service
PAF	postcode address file
PIN	personal identity number
RDD	random digit dial
STD	sexually transmitted disease
STI	sexually transmitted infection
T-ACASI	telephone-audio computer assisted self interview
UAI	unprotected anal intercourse
URL	universal resource locator
WWW	World Wide Web

PART I

INTRODUCTION

1 Introduction

1.1 Overview of the thesis

The aim of this study is to examine the issues surrounding the recent and rapid emergence of the Internet as a new research tool, in order to inform those wishing to undertake research via the Internet. Use of the Internet is often recommended for its ability to reach minority populations that are otherwise difficult to access and its capacity to facilitate reporting sensitive behaviour. This has no doubt given rise to an increased interest in Internet methods for the investigation of HIV risk behaviours and potential preventive measures. Yet much remains to be learnt about the impact of the Internet on accessing minority populations and gathering sensitive data. This thesis sets out to address these two broad issues in the context of online research into sexual behaviour among gay and bisexual men.

The thesis takes an empirical approach to the exploration of the methodological issues surrounding the use of the Internet for quantitative and qualitative data collection. Part one establishes the background to the research. **Chapter 1** begins this account with an examination of the three key elements of the study: the emergence of the Internet, the population of gay and bisexual men and data of a sexual nature. **Chapter 2** goes on to describe the context of the thesis: the Internet and HIV study, a Medical Research Council funded project, which set out to examine sexual risk behaviour among gay and bisexual men. **Chapter 3** concludes the background to the research with an overview of the methodology and analytical strategy for the thesis.

Parts two and three describe the empirical research undertaken. Part two (Chapters 4, 5 and 6) addresses the issue of use of the Internet for capturing samples of gay and bisexual men and part three (Chapters 7 and 8) examines the impact of the Internet mode on reporting in studies of sexual behaviour. Because each of these chapters involves a different methodological research question which is addressed by discrete

analysis, each chapter opens with a brief synopsis of the method and findings and contains a review of the relevant literature and description of the methods, before going on to describe and discuss the findings. The research questions and findings from these chapters can be summarised as follows:

Chapter 4

Why do gay and bisexual men participate in web surveys of sexual behaviour?

The opportunity to contribute to research was the most important reason for taking part in the web survey, followed by the topic of the survey. The findings indicate that the altruistic desire to contribute to research into sexual health promotion and HIV prevention was one of the principal reasons for participation.

Chapter 5

Among gay and bisexual men, who drops out of web surveys of sexual behaviour, at what point and why?

Drop out was much more likely to occur in the very early stages of the survey. Drop out in the later stages tended to be more likely among men who were younger, minority ethnic and not open about their sexual orientation. Respondents indicated that lack of time was the most important reason for quitting.

Chapter 6

How do self-selected Internet samples of gay and bisexual men differ from men who have sex with men drawn from a probability sample?

The national Internet samples of gay and bisexual men were broadly similar to the sub-sample of men who have sex with men (MSM) drawn from a national probability sample on a range of sociodemographic and behavioural variables with the exception of those indicating HIV risk behaviour. Differences were more pronounced when comparing the London samples, particularly for questions about sexual risk behaviour.

Chapter 7

What is the modal difference in item nonresponse when gay and bisexual men complete web and pen-and-paper surveys of sexual behaviour?

Use of the web survey reduced item nonresponse to a considerable degree. It did not appear to reduce item nonresponse to potentially sensitive questions over and above the general effect of mode on all questions.

Chapter 8

What is the modal difference in interview experience when gay and bisexual men are interviewed about sexual behaviour face-to-face or by online synchronous chat?

The use of online synchronous interviews increased the speed of turnaround. However, the quantitative-qualitative hybrid nature of the online interview meant that some respondents were unsure about how to respond in this context. The medium of chat was less well suited to the emotional and intellectual demands of the interview, given the additional burden of translating thoughts into typing.

Chapter 9 forms the fourth and final part of thesis. It draws together the findings from parts two and three to discuss the use of the Internet in research into sexual behaviour among gay and bisexual men.

1.2 Foundation elements

Three key foundation elements are involved in this examination of use of the Internet for collecting data on sexual behaviour among gay and bisexual men. These are the mode of data collection, the target population and the nature of the data. This chapter aims to place each of these elements in context, in order to lay the foundations for the methodological research which follows. It provides a brief history of the use of technology for data collection in social research, the emergence of the Internet and its adoption as a data collection tool. It goes on to describe the complexities involved in researching the population of gay and bisexual men.

Finally, it examines the issues associated with research involving data of a sexual nature and the concept of sensitivity that is generally applied to it.

1.2.1 Emergence of the Internet as a data collection tool

1.2.1.1 Use of technology for data collection

In order to locate the emergence of the Internet as a data collection tool, the following begins with an account of how technologies other than the Internet have been used for data collection in social research. It is concerned with data collection that is mediated by technology in the same way that this study is concerned with data collection that was mediated by the Internet. The two methods of interest here are survey research and qualitative interviewing.

Despite the vast literature on the use of technology in survey data collection, there is relatively little in the literature on technology-mediated qualitative interviewing. The lack of such literature is likely to be because qualitative interviews are almost invariably conducted face-to-face (Legard *et al.* 2003). Although there is a place for the telephone in qualitative interviewing (Burke & Miller 2001), the general preference for the face-to-face mode means that the equipment used to *record* such interviews has been the main application of technology in the data collection process. Such equipment has certainly become cheaper, better, more acceptable and more widely used (Gibbs *et al.* 2002) but there appears to be no widespread chronological development in the use of new technology in qualitative interviewing. Indeed, the Internet is the only use of technology described by Fontana and Frey (2000) in their account of qualitative interviewing.

In the light of this, the what follows will focus on the expanding use of technology in the survey data collection process and qualitative interviewing will be re-introduced subsequently in the description of use of the Internet for data collection. The following provides a brief history of the adoption of technology for survey data collection which illustrates how the Internet can be situated in a chronology of technology-mediated survey methods. It relies to a great extent on a monograph by Couper and Nicholls (1998), *"The history and development of computer assisted*

survey information collection methods" for the historical details of the computerisation of the survey process and the reader is referred to this paper for a fuller treatment of the subject.

Although the telephone was the first piece of technology to be used in survey data collection, it was not until comparatively recently that this mode was widely adopted. In 1936, the Literary Digest famously predicted that Landon would beat Roosevelt in the presidential election based on their biased sample of car owners and telephone subscribers (Webb & Wybrow 1974). This mistaken forecast can be attributed to the fact that only 35% of US households had telephones in 1936. It took nearly forty years and a dramatic increase in telephone coverage for US survey researchers to accept the legitimacy of such surveys (Massey 1988). US coverage was still only around 80% by the beginning of the 1960s but the following years saw an increase to its current level of 96% (Tourangeau 2004).

The acceptance of telephone interviewing was also no doubt assisted by additional new technology which facilitated its use. The first nationwide telephone interview facility was established in the United States in 1966. This was followed by the first computer assisted technique with the invention of computer assisted telephone interviewing or CATI which was first implemented in the US in 1971 by Chilton Research Services for a survey by AT & T to evaluate customer satisfaction with their telephone services (Nicholls & Groves 1986). Despite the rapid escalation in the use of this technology in the United States, it was not adopted in Europe until the 1980s. Although a study by Collins *et al.* (1998), reports a dramatic increase in the adoption of CATI by UK market research companies over the 1980s and 1990s, many such companies continued to operate without it and face-to-face interview modes have continued to dominate in Europe (de Leeuw & Nicholls 1996; Jeavons 1998).

Whereas CATI was first used in the US commercial sector, European government statistical agencies were the first to implement computer assisted personal interviewing or CAPI. This is attributed to the commitment of these agencies to

face-to-face methods and their ability to afford the equipment (Couper & Nicholls 1998). Early efforts to develop appropriate portable computer technology were superseded by the commercial launch of laptops in the mid-1980s which was followed by the first use of CAPI in a national survey by the Dutch Labour Force Survey in 1987. CAPI was first used in a British government survey in 1990 and most of such surveys had converted to CAPI by the mid-1990s (Martin & Manners 1995). Its use also quadrupled among UK market research companies over the 1990s (Collins *et al.* 1998).

Whereas all leading European and North American governmental and private sector survey organisations had moved or were moving over to computerisation of interview surveys by the end of the 1990s, the take up of computerised alternatives to postal surveys was slower and mainly restricted to North America. Although self-completion computerised questionnaires were first used in the early 1980s, the introduction of computer technology has therefore had the least impact on postal surveys (Couper & Nicholls 1998; Tourangeau 2004). One of the earlier methods of this type was disk-by-mail (DBM), where a floppy disk containing an electronic version of the survey was sent by post to the respondent, and another was the use of electronic mail surveys (Couper & Nicholls 1998). Tourangeau (2004) describes web surveys as the third generation of computerised counterparts to postal surveys, following DBM and email surveys.

Another type of computerised self-completion is computer assisted self interview or CASI. This may form part of a CAPI interview when the computer is turned towards the respondent so that they can read and enter their own responses instead of relying on the interviewer. A variation on this is audio-CASI where respondents hear the questions over a headset while reading them from the computer and then key their answers into the computer. The telephone version is T-ACASI where respondents hear the questions over the phone and respond using the telephone keypad.

1.2.1.2 Emergence of the Internet

The Internet began to emerge as a potential data collection tool in the 1990s. What follows is a description of the historical development of the Internet which draws mainly from information provided in two online resources which are both entitled *"A brief history of the Internet"* by Howe (2004) and Leiner *et al.* (2003), before turning to the use of the Internet for data collection.

As these sources recount, the earliest carnation of the Internet was ARPANET which came online in 1969 when research led by the Advanced Research Projects Agency enabled connection of computers in four US universities. As further institutions were gradually added to the system, ARPANET was adopted by computer experts, engineers, scientists and librarians who needed the technical ability to master the complicated system. Following the first public demonstration of ARPANET in 1972, the subsequent years saw a number of developments before the emergence of the Internet in the form that we would recognise today. Thus, email was adapted for ARPANET in 1972 to become the largest network tool for over a decade (Leiner *et al.* 2003). The first international connections were made when the UK and Norway joined the network in 1973 (Hewson *et al.* 2003). USENET, which is not considered as part of the Internet, began in 1979 and the newsgroups that it facilitated allowed shared access to the communication forums. The first MUD (Multi-User Dungeon) was produced in 1979, enabling participants to develop online personae to interact with each other. 1981 saw the development of BITNET for which Listserv was developed allowing further communication possibilities through use of this and other email discussion lists. JANET, the Joint Academic Network, was established in 1984 in the UK (Hewson *et al.* 2003). Internet relay chat or IRC began in Finland in 1988 and took off globally in 1991 when it was used to get up-to-date information on Iraq's invasion of Kuwait after radio and television broadcasts had been cut off (Living Internet 2005).

The World Wide Web was invented by Tim Berners-Lee, a scientist at CERN (L'Organisation Européenne pour la Recherche Nucléaire) in 1989. A key feature of the web is that it is based on hypertext, a system whereby links embedded in one text

enable connection to other texts. The web is, however, just one of many Internet-based services and is a distinct entity from the Internet. Thus, the Internet has been likened to a road system which requires the three essential elements of physical conditions, common behaviour and services of which the web is one (CERN 2005). The physical conditions are the cables and hardware which link computers and can be represented by a road map. The common behaviour is the Internet protocol which provides the rules whereby computers are able to exchange data in the same way that traffic rules govern effective use of the road system. And just as services such as public buses and delivery vans use the road system, so services such as email and the web use the Internet.

1993 heralded the launch of Mosaic, the first point-and-click graphical browser, from which Netscape evolved to become the most successful tool until the release of MicroSoft's Internet Explorer. CERN (2005) name 1994 "Year of the Web" because it saw the world's first international World Wide Web conference and a surge in web use that is illustrated by its 10,000 servers and 10 million users by the end of the year. It also saw the release of HTML forms which enabled users to communicate back to server and was a necessary development for the implementation of web surveys (Kehoe & Pitkow 1996). The subsequent release of the programming language Java signalled a further leap forward in the retrieval, display and use of information over the Internet (Musch & Reips 2000).

Although the Internet was originally funded by government and limited to research, educational and governmental use, restrictions on commercial use officially ended in May 1995 with the discontinuation of the National Science Foundation's sponsorship of the Internet backbone. AOL, Prodigy and CompuServe were the first commercial networks to come online and the entry of MicroSoft completed the move to a commercially based Internet. By this time, the Internet was no longer beyond the technical ability of the average person (Houston & Fiore 1998) and the scene was set for its widescale adoption.

The early image of the Internet user was a young, white male with high education and socio-economic status but the profile of Internet users has become more diverse over the years. Despite the dramatic increase in use of the Internet, however, certain disparities remain. Thus, British national statistics show that the proportion of males who had recently used the Internet increased from 47% to 64% between 2000 and 2004 whereas the proportion of females increased from 33% to 58%, and the proportion of 16-24 year olds increased from 70% to 90% whereas the proportion of 55-64 year olds increased from 24% to 48% (Office for National Statistics Omnibus Survey, http://www.statistics.gov.uk). Similarly, the lowest income households with Internet access increased from 3% to 15% between 1998-99 and 2003-04, whereas the highest income households saw a considerable increase from 32% to 89% (Office for National Statistics Family Expenditure and Expenditure & Food Surveys, http://www.statistics.gov.uk). The Oxford Internet Survey also found a continuing disparity in the education of current Internet users in 2005, with 27% of people with no qualification using the Internet compared to 66% of those with school leaving qualifications and 83% of those with a degree level qualification (Dutton *et al.* 2005).

1.2.1.3 Use of the Internet for data collection

Despite the above differences in Internet take up, survey researchers have embraced the potential of the Internet, describing it as,

> *"the hottest, fastest growing and most revolutionary development to change our industry since George Gallup demonstrated in 1936 that 'scientific' opinion polls were the best way to predict elections"* (Taylor 2000: 51).

There is little doubt that the Internet is having an enormous impact on survey research. For the time being, it seems that qualitative interviewers are unlikely to adopt the Internet on anything like the same scale, but it remains a revolutionary development with future possibilities that few can ignore.

At the present time, many questions remain about the potential and credibility of Internet data collection and this thesis sets out to examine some of these issues in detail. Agreement has, however, emerged on the general benefits of Internet

research. It is fast and economical, has the potential to access distant and minority populations and facilitates the collection and collation of data. These advantages apply to a range of Internet methods whereas its additional capacity to reach large samples is of particular importance to survey research. Probably the main disadvantage at the present time is that Internet use remains far from universal.

The Internet can be used to collect data in many ways, from ethnographic studies of online communities to tracking usage through analysis of log files. The present study, however, is concerned with data collected in surveys and qualitative interviews under conditions where respondents are explicitly offered the opportunity to participate. The following will briefly describe how the Internet developments described earlier have been incorporated into these forms of data collection. A more detailed discussion of these methods is given in parts two and three, by way of rationalisation for the empirical research undertaken here.

Online surveys

Online surveys can be divided into two main forms – the email survey and the web survey. Email was one of the earliest forms of Internet-based communication and was indeed used in online self-completion surveys before the introduction of the Internet as we know it (Batagelj & Vehovar 1998). Thus, Kiesler and Sproull published the first article on email surveys, *"Response effects in the electronic survey"*, in 1986. Email surveys can be embedded in the text or sent as attachments. Broadly speaking, the respondent can complete and return them electronically or print them off to complete by hand and return by fax, post or any other available method. They are technically simple and non-interactive and have been described as online pen-and-paper surveys (MacElroy 1999).

Although there is quite an extensive literature on email surveys, the development of web surveys is likely to eclipse their use. Whereas earlier computer based methods had limited impact on postal surveys (Couper & Nicholls 1998; Tournageau 2004), the introduction of web surveys is said to,

> *"represent a huge increase in the use of self-completion surveys and is a radical change in direction for survey data collection"* (Jeavons 1998: 1).

Dillman and Bowker (2001) argue that web technology caught survey methodologists by surprise and that the impetus to develop such surveys came mainly from computer programmers. Thus, the first publicly accessible web survey was developed by computer scientists at the Graphics, Visualization and Usability Center (GVU) of the Georgia Institute of Technology and posted on the Internet in January 1994, in order to explore the profile of web users, their reasons for using the web and their opinions about it (Kehoe & Pitkow 1996). It achieved nearly five thousand responses over a one-month period (Pitkow & Recker 1994). Over the decade that has passed since the first GVU survey, innumerable web surveys have been used to recruit countless respondents in a range of different ways, as Chapter 6 will illustrate. Although web survey methodology has made great advances, the technology continues to evolve and it is important that survey researchers continue to work closely with the programmers to bring the expertise of both parties to bear on web survey design (Dillman & Bowker 2001).

Online interviews

Market research organisations were among the first to realise the potential of the Internet for qualitative research (Chase & Alvarez 2000). Thus, moderated email groups (MEGs) were developed in this sector whereby respondents individually supply answers to questions by email and these are summarised by a group moderator and sent back for comment (Adriaenssens & Cadman 1999; Eke & Comley 1999). Other researchers have pioneered the use of online interviews by simply entering into an email correspondence with individual respondents. Whether involving groups or individuals, this type of interview is known *asynchronous* because the communication between the interviewer and respondents is not instantaneous. It is thus distinguished from interviews that are *synchronous* and take place in "real time", involving communication that is more or less immediate. Synchronous interviews can be carried out using any online instant forms of communication, such as Multi User Dungeons (MUDs) or IRC, and they can be conducted with groups or one-to-one. Although qualitative researchers have not

rushed to use online interviews, there is a growing literature in this area that is described in Chapter 8 which also provides an empirical analysis of the use of online synchronous interviews for collecting data on sexual behaviour.

1.2.2 Researching the population of gay and bisexual men

The above has summarised how the Internet emerged as a data collection tool. Part two of this thesis will go on to examine its potential for conducting research among gay and bisexual men. What follows is an introduction of the issues associated with researching this population which also aims to clarify some of the terminology that will be used in this thesis.

Defining a target population of gay and bisexual men is a complex matter. Whatever criteria are adopted, however, the fact remains that these men form a tiny minority of the general population. In 2000, just 2.8% of British men reported sex with another man in the previous 5 years (Mercer *et al.* 2004) which means that finding effective methods of reaching this group is a significant challenge and a key concern of this thesis.

Research with this group is often referred to as involving a *hard to reach population.* Pickering (2003) describes such populations as falling into two categories. The first is minority groups which are included on standard sampling frames but are such a small proportion of the population that standard probability sampling becomes problematic. Examples are minority ethnic groups or gay and bisexual men. The second is special groups, which are not included on standard sampling frames, and may refer to groups such as homeless people. It is argued that the social stigma that continues to be associated with homosexuality creates a *hidden population* of gay and bisexual men whose members are not easily identified (Kalton 1993). However, this population does not form a homogenous group (Sandfort 1997) and the use of the term hidden population would be inappropriate for men who are open about their sexual orientation.

In the literature, the term "gay" is generally applied to the group of men who self-identify as gay (Sell & Petrulio 1996). It is used here alongside "bisexual" to describe the men who participated in the Internet and HIV study (described in Chapter 2) because the vast majority of these men self-identified as gay or bisexual. Such self-identification was indeed one of the criteria for eligibility in parts of the study. The problem of how to define a population of "gay and bisexual" men is, however, infinitely complex. Sexual behaviour is just one of a range of dimensions that are used to define homosexuality. These dimensions, which include feelings, physical or emotional attraction, fantasies and self-definition, do not necessarily overlap and are subject to change (Sandfort 1997). As Plummer (1981) has written,

> *"there is, in short, no absolute 'fit' or congruency between doing, thinking or feeling, and there is no necessary fit between any of these and the act of labelling oneself as 'a homosexual'"* (Plummer 1981: 67).

Thus, an online survey of self-identified lesbian, gay, bisexual and transgender (LGBT) college students found that although 90% of males were primarily or exclusively attracted to males, 28% had engaged in penile-vaginal intercourse over their lifetime and 24% had performed oral sex on a female (Lindley *et al.* 2003). Although all respondents self-identified as LGBT, males and females who were at least 25 years old were significantly more likely to report sex with a same sex partner than younger respondents. Boulton and Fitzpatrick (1996) describe a typology of bisexual men which includes men who identify as gay and their emphasis on the fluidity and diversity of bisexuality furthermore illustrates the problem of definition. In the light of such complexity, Martin and Dean (1993) argue that it is unlikely that experts would ever agree on the boundaries of homosexuality that are the essential precursor to drawing up a sampling frame of homosexual males.

This lack of consensus is illustrated in the work of earlier eminent researchers in the field of sex and sexuality. Kinsey and his colleagues conceived a homosexual-heterosexual rating scale which combined sexual behaviour and psychosexual reactions on a continuum from zero which is fully heterosexual to six which is fully homosexual (Kinsey *et al.* 1948). Although the scale includes the psychosexual,

Kinsey's work tended to emphasise sexual behaviour whereas Masters and Johnson emphasised the psychosocial aspects. They defined lesbians and gay men as,

> *"men or women who have a preferential sexual attraction to people of their same sex over a significant period of time"* (Masters *et al.* 1995: 372).

Similarly, bisexuals were described as men and women who are attracted to either sex. Although homosexuals and bisexuals were likely to engage in same sex sexual activity, this was not considered a necessary condition for the definition of a homosexual or bisexual person.

Given the problem of definition, it is not surprising that a review of the literature on public health research in the early nineteen-nineties involving male and female homosexuals found that only 4 of 152 studies (2.6%) conceptually defined the population they were sampling (Sell & Petrulio 1996). Although most of the studies described the operational method used to identify respondents, with self-report and setting the most commonly used, the researchers concluded that it was impossible for the studies that they identified to generalize to a population of homosexuals without including important caveats. In line with this, Sandfort (1997) argues that we know little about the homosexuality of men who are not connected to the gay community. He points out that Humphreys' (1970) ethically controversial work *"Tearoom Trade: A Study of Impersonal Sex in Public Places"*, which used covert methods to explore male sexual behaviour in public conveniences, was an unusual study that did not sample directly from the population of men who participate in gay culture.

In order to include men who are not acculturated into the gay community, an alternative expression that is often used in the research is "men who have sex with men" or MSM. The use of MSM may, indeed, seem more appropriate in surveys of the type used by the Internet and HIV study which are not so much concerned with sexual identification as with sexual behaviour. In such surveys the target population may be defined as men of a particular age, living in a particular place who have been sexually active with at least one other man over a defined period of time. The terms "gay" and "bisexual" are used extensively throughout this thesis, however, because

much of the research to which it refers is based on samples of men who self-identify in this way. Furthermore, the use of "MSM" has been associated with making gay men less visible and undermining the gay movement (Sandfort 1997).

There are circumstances, however, where MSM is the most appropriate term to use. For example, this thesis involves analysis of a sub-sample of MSM drawn from the British National Survey of Sexual Attitudes and Lifestyles (Natsal) 2000 (as detailed in Chapter 6). This sub-sample was defined according to the criteria described in the previous paragraph and because Natsal did not contain any questions about sexual identification, no assumptions can be made about the sexual identification of these men who report that they have been sexually active with men. Men in the Natsal sample are therefore described throughout this thesis as MSM.

This thesis is therefore concerned with both the population of self-identified gay and bisexual men and the population of MSM. As Figure 1.1 illustrates, there is a substantial overlap between these populations but this overlap does not account for men who self-identify as gay or bisexual yet do not have sex with men or men who have sex with men but do not identify as gay or bisexual.

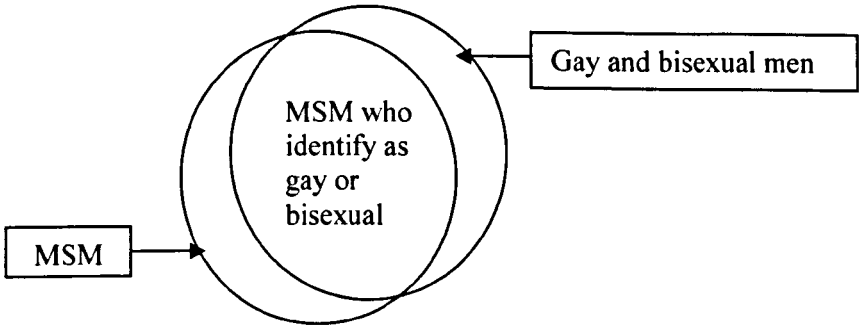


Figure 1.1: Populations of MSM and self-identified gay and bisexual men

This latter category is likely to be missed in traditional sexual health research because, as Chapter 6 describes, men are generally recruited from gay community

contacts and / or sexual health services. The Internet may, however, offer an important new site for sampling MSM which part two of this thesis will consider.

Even studies that define the target population based on sexual behaviour must be specific about the nature and timescale of the same sex activity that leads to inclusion. Different definitions of same sex activity will lead to different estimations of prevalence as illustrated in Table 1.1. It uses data from Natsal 2000 and its predecessor Natsal 1990 to show how different questions elicit different proportions of same sex activity within the male population.

Table 1.1: Response to questions of sexual behaviour

	Natsal I (%) 1990	Natsal II (%) 2000
Ever had sexual experience, not necessarily including genital contact, with a partner of the same sex?	5.3	8.4
Ever had sex with a same sex partner, including genital contact?	3.7	6.3
Have you had a same sex partner in the last five years?	1.5	2.6

Source: National survey of sexual attitudes and lifestyles II: Reference tables and summary report (Erens et al. 2003)

The table also shows how responses to these questions have changed between 1990 and 2000. Such changes are attributed to a true change in prevalence but also to a greater willingness to report same sex activity, due to improved survey methodology and more tolerant social attitudes (Johnson *et al.* 2001b). This illustrates how even when the researcher has a clearly defined inclusion criteria, the achieved sample may depend to a certain extent on respondent willingness to report certain sexual behaviours. Although attitudes are changing, with data from Natsal 2000 showing that 88.7% of gay and bisexual men considered that homosexual sex was "not wrong

at all", compared to 74.1% in 1990 (Mercer *et al.* 2004), disclosure of homosexual behaviour undoubtedly remains an issue for some men. Such disclosure may also be affected by the survey mode, with 7.1% of women in Natsal 2000 reporting a same-sex experience in the face-to-face interview compared to 9.7% in the self-completion mode (McManus 2003).

The above summary has introduced some of the issues that arise in researching sexual behaviour among gay and bisexual men. It shows that the samples achieved are dependent on how the target population is defined and how the respondents define themselves and it illustrates the problems of generalising from such research. It thereby sets the scene for part two of this thesis which examines the role of the Internet in sampling this population. This assessment of the Internet is undertaken by addressing the following questions, each of which will be fully elaborated in Chapters 4, 5 and 6:

- Why do gay and bisexual men participate in web surveys of sexual behaviour?
- Among gay and bisexual men, who drops out of web surveys of sexual behaviour, at what point and why?
- How do self-selected Internet samples of gay and bisexual men differ from men who have sex with men drawn from a probability sample?

1.2.3 Collecting data on sexual behaviour

Part three of this thesis is an examination of use of the Internet mode in studies of sexual behaviour. The topic of such studies is often described as sensitive and the following sets out to examine the application of this concept to the Internet and HIV study, beginning with a discussion of what is meant by sensitive data.

A meta-analysis of studies of socially desirable responding coded questions as sensitive if they asked for what the authors described as, *"personal information not normally discussed among casual acquaintances"* such as matters of finance,

criminal behaviour, sexual behaviour and drug use (Richman *et al.* 1999). Throughout the literature, these topics are described as sensitive without further elaboration, lending support to Lee and Renzetti's (1990) assertion that the label "sensitive" is often applied as if it were self-explanatory. They argue that any topic may potentially be sensitive, depending on its context, and provide the following broad definition,

> *"a sensitive topic is one which potentially poses for those involved a substantial threat, the emergence of which renders problematic for the researcher and/or the researched the collection, holding, and/or dissemination of research data"* (Lee & Renzetti 1990: 512).

Tourangeau and Smith (1996) also provide a definition which describes sensitive questions as those which raise concern about disapproval or other negative consequences, or are otherwise considered as an invasion of privacy. These definitions recognise that question sensitivity is not an invariable property of the topic but something that Catania (1997) argues will vary as a function of factors including respondent and interviewer characteristics, the recency of the behaviour in question and the wording of the question.

Respondent perception of the question is therefore of key importance in determining its sensitivity to that individual and how such perceptions are shaped may be understood with reference to the concept of social desirability. Groves (1989) describes social desirability in the following way:

> *" 'Social desirability' is a judgement about how highly valued a particular attribute is in the culture. Some attributes of individuals are negatively valued by societal norms ... Other attributes are positively valued ... The concept requires an evaluation by a person concerning what the society judges as a desirable trait. These judgements need not be consistent over persons in the same group"* (Groves 1989: 437).

One of the manifestations of judging something to be socially undesirable may be item nonresponse which has been found to increase with increasing levels of threat or sensitivity (Sudman & Bradburn 1973; Tourangeau *et al.* 2000). More often,

however, the impact of social desirability has been examined in terms of its effect on self-reports of sensitive behaviours. Since people want to present a positive image of themselves (Catania *et al.* 1990), social desirability bias is said to occur,

> *"if respondents perceive their behaviour as being deviant from a socially acceptable norm and modify their responses either by understatement or overstatement to bring them closer to that norm"* (Wadsworth *et al.* 1996: 112).

Socially desirable responding has generally been captured by one of two measures. The first is the Marlowe-Crowne scale (Crowne & Marlowe 1960) which was developed to measure the need for social approval as a personality trait. The second is the Balanced Inventory of Desirable Responding (BIDR) which was developed by Paulhus (1991) and has two sub-scales. The first measures impression management or the deliberate manipulation of self-report in specific situations and the second measures self-deceptive enhancement whereby respondents actually believe their positive self-reports.

The use of such scales provides an objective measure of socially desirable responding but an understanding of how social desirability shapes answers to sensitive questions requires an appreciation of what is socially acceptable to the respondent. Groves (1989) points out that this need not be consistent between members of the same group which may be partially attributed to the multiplicity of cultural norms to which individuals may refer (Catania 1997). The empirical evidence, however, indicates that group differences may be particularly powerful for certain questions of sexual behaviour.

In studies of sexual behaviour, men are found to report higher mean numbers of sexual partners and women to report lower numbers than would be possible in a closed population (eg Wellings *et al.* 1994). The social norm whereby having multiple sexual partners is more likely to be applauded in males and censured in females is generally given by way of explanation for these findings. Although the discrepancy can be reduced when certain assumptions are made, it is not entirely removed. Thus, Wadsworth *et al.* (1996) reduced the disparity by accounting for the

age mixing in male-female sexual partnerships, the under-representation of female sex workers and lower male response rates. The gap was also narrowed from 3.2:1 to 1.2:1 when respondents with greater than 20 lifetime sexual partners were excluded from the analysis (Morris 1993).

Another examination of male and female differences in reporting sexual behaviour found that male college students were likely to positively bias their responses in questions of condom use and attitude towards condoms whereas females were not likely to do so (Agnew & Loving 1998). The bias in male reporting was somewhat reduced where instructions emphasised respondent anonymity.

In the same way that these gender differences have been found in reporting particular sexual behaviours, other group differences are likely to exist for other questions of sexual behaviour. A good example is that of sexual orientation and anal sex which is arguably one of the cases where,

> *"what is contra-normative behavior in the society as a whole may be socially valued by some subcultures and therefore overreported by its members"* (Bradburn *et al.* 1978: 222).

Whereas anal sex is considered to be a stigmatised activity within certain sections of the heterosexual population, only *unprotected* anal sex with a non-primary partner is similarly stigmatised among the gay male population, for whom anal sex with a condom may be considered socially desirable (Acree *et al.* 1999). In an attempt to reduce the social desirability bias associated with reporting unprotected anal sex, Difranceisco *et al.* (1998) re-framed survey questions to ask about total amount of anal sex followed by frequency of condom use. These questions were associated with higher reporting of unprotected anal sex than questions which asked separately about sex with and without a condom, particularly for men who had more frequent anal sex.

The distinction in the social acceptability of anal sex illustrates the "range of sensitivities" associated with questions of sexual behaviour (Catania 1997). It

suggests that the reliable reporting of one sexual behaviour does not necessarily entail the reliable reporting of another (Catania *et al.* 1990; 1993). The example also highlights the impact of the broader societal context on shaping socially acceptable behavioural norms. Gay men nowadays may be more reluctant to report casual unprotected anal sex, whereas prior to the AIDS epidemic, it may have been more readily reported (Catania *et al.* 1990).

The differences in the social acceptability of certain sexual behaviours suggest that there may be less consensus among the population as a whole on the extent to which they find other topics to be sensitive than has been assumed. Thus, it is unsurprising that Tourangeau *et al.* (2003) found a low level of internal consistency (Cronbach's $\alpha = 0.19$) for eight questions measuring healthiness of lifestyle, use of recreational drugs, voting and church attendance that were purposively selected as sensitive topics. An earlier study by Tourangeau and his colleagues (1997b) also highlights the variable nature of topic sensitivity. When respondents were led to believe that the truth of their answers to a range of sensitive questions was detectable by a physiological recording device, they were generally less likely to give more socially desirable answers. The magnitude and significance of these differences, however, was found to vary both between and within sets of questions on drinking, sexual behaviour and recreational drug use.

Such differences in reporting may also be related to the individual's style of self-presentation. Whereas some people have a general tendency to present a socially acceptable self, others vary their self-presentation according to the situation (DeMaio 1984). People may therefore vary in the degree to which their reports of sexual activities are subject to social desirability (Catania *et al.* 1990).

The relationship between the respondent's personal experiences and the substantive content of the question is also of key importance. Thus, Copas *et al.* (1997) found that embarrassment and refusal to complete the more sensitive questions in the self-administered questions of a survey of sexual behaviour were associated with less varied sexual experience. An association between item nonresponse to such

questions and less sexual experience was also found in the US General Social Survey (Wiederman 1993).

Although the evidence indicates that question sensitivity is not a dichotomous attribute, Groves (1989) argues that much of the research on socially desirable responding does not address this issue. He finds that,

> *"researchers, based on their own judgements, have labelled attributes as socially desirable or undesirable and interpreted responses to the questions using those labels. Such an approach is wise only when there is homogeneity in the population on the extent of social desirability of the attribute"* (Groves 1989: 447).

As described earlier, the population of gay and bisexual men is by no means a homogenous group and, as Groves suggests, this may result in a lack of consensus as to what constitutes socially desirable behaviour. In the light of this potential diversity, part three of this thesis focuses on the impact of the Internet mode in studies of sexual behaviour. It thus seeks to address the following questions which will be further developed in Chapters 7 and 8:

- What is the modal difference in item nonresponse when gay and bisexual men complete web and pen-and-paper surveys of sexual behaviour?

- What is the modal difference in interview experience when gay and bisexual men are interviewed face-to-face or using online synchronous chat?

This chapter has sought to lay down the foundation elements involved in this thesis: the emergence of the Internet, the population of gay and bisexual men and data of a sexual nature. These elements were integral to the context within which the methodological research for this thesis took place, the Internet and HIV study. To illustrate this further, Chapter 2 gives a summary of the Internet and HIV study and the methods that it used. It provides the background for Chapter 3 which is a detailed description of the research strategy adopted in this thesis and its relationship to the Internet and HIV study.

2 Internet and HIV study: the context of the methodological research

2.1 Introduction

The Internet and HIV study provided the context for this research into the use of the Internet as a data collection tool. The methodological research undertaken for the thesis involved secondary analysis of data collected for the Internet and HIV study and collection of additional primary data from respondents who took part in the Internet and HIV study. This chapter summarises the design and methods of the Internet and HIV study in order to provide the framework for the chapters which follow and Chapter 3 goes on to describe the research strategy that was adopted.

A detailed description of the design and methods employed by the Internet and HIV study can be found in *"The Internet and HIV study: design and methods"* by Elford *et al.* (2004a) which is available at: http://www.biomedcentral.com/1471-2458/4/39 . The main findings are described in *"Gay men who look for sex on the Internet: is there more HIV/STI risk with online partners?"* by Bolding *et al.* (2005). They indicate a tendency for HIV positive gay men in London to meet casual partners of the same HIV status for unprotected anal intercourse (hereafter referred to as UAI) through the Internet. This does not present a risk of HIV transmission because both partners are HIV positive but it does present a risk for STI transmission. Although UAI with a partner of unknown or discordant HIV status presents a risk for HIV transmission, gay and bisexual men were not found to be more likely to meet casual UAI partners of unknown or discordant HIV status online rather than offline. Taken together these findings suggest that the Internet does indeed contribute to STI transmission, particularly among HIV positive men, but there was no evidence that the Internet *per se* creates a risk for HIV transmission (Elford 2006).

The Internet and HIV study was funded by the UK Medical Research Council and Department of Health under the AIDS Epidemiological Research Programme

(G0100159). A number of earlier studies had found an association between seeking sex on the Internet and high risk sexual behaviour. The Internet and HIV study therefore set out to investigate whether the Internet represents a new sexual risk environment for gay and bisexual men living in London and, if so, to examine its underlying processes. It also sought to consider the potential of using the Internet for HIV prevention. It focused on London men because of the higher incidence and prevalence of HIV infection among gay men in the capital.

The Internet and HIV study employed both quantitative and qualitative methods in order to draw a fuller picture of how gay and bisexual men use the Internet for meeting sexual partners. What follows begins with a description of the quantitative arm and goes on to describe the qualitative arm. It concludes with a description of the ethical protocol employed in the Internet and HIV study which also provided the framework for the ethical approach adopted in this methodological study.

2.2 Internet and HIV study: quantitative methods

The Internet and HIV study recruited four samples of gay and bisexual men for the quantitative arm, with the aim of triangulating data from a number of sources and gathering samples with relevant features to address the research question. The following describes each of these four samples (Internet, gym, HIV positive clinic and HIV testing clinic) in order to provide a full description of the Internet and HIV methods. However, only data from the Internet and gym samples were used in the methodological research which forms the basis of this thesis. The gym, HIV positive clinic and HIV testing clinic samples completed pen-and-paper questionnaires whereas the Internet sample completed a web survey.

Internet sample

Men who use the Internet to seek sexual partners were of key importance to the Internet and HIV study. In 2002 and again in 2003, men who used UK chat rooms and personal profiles on gaydar (http://www.gaydar.co.uk) and gay.com (http://UK.gay.com) were invited to participate in the survey. These are two of the most popular websites for gay men in the UK. In May and June each year, men

could click on pop-ups and banners that appeared on these websites and took them directly to the homepage of the web survey. Although the target population was London men, it was not technically possible to restrict the pop-ups and banners to London chat rooms and profiles only. The pop-ups and banners therefore appeared in UK chat rooms and profiles. Although these are used mostly by men in the UK, some non-UK residents also use them. The surveys drew worldwide samples of 4,974 men in 2002 and 2,752 men in 2003. It is estimated that these samples represent less than 1% of the people using the chat rooms and profiles over the survey period. Whereas only data from London men were used in the Internet and HIV study, the methodological research undertaken for this thesis includes data from the Internet samples at the London, national (British and UK) and worldwide level.

Gym sample

Gay and bisexual men who use central London gyms are broadly representative of London gay men "on the scene", who go to gay bars, clubs and so on (Elford 2002). Gym-based surveys are therefore an effective method of gathering detailed data on sexual behaviour from this group of men (Elford *et al.* 2004a). In 2002 and again in 2003, all men who used one of seven central London gyms during a one-week period between January and March were invited to participate in the survey. All seven gyms had a substantial gay membership (40-100% of clients) but only men who identified as gay or bisexual were asked to answer the questions on sexual behaviour and the Internet. Completed pen-and-paper questionnaires were returned to collection boxes in the gym or posted back to the research team. In 2002, the gym sample consisted of 914 gay and bisexual men who completed the survey and provided information on their HIV status and in 2003, it consisted of 543 men.

HIV positive clinic sample

HIV positive men were over-sampled to allow for in-depth study of their sexual risk behaviour and Internet use. Gay and bisexual men attending a National Health Service (NHS) outpatient HIV clinic are considered to be broadly representative of those living with diagnosed HIV. From October 2002 to May 2003, HIV positive men attending an NHS outpatient treatments clinic were invited to participate in the

survey. Over the survey period, 1,001 men attended the clinic of whom 939 were considered eligible for the study on the basis of language and health. It was not possible to make contact with everyone who passed through the clinic, but of the 864 men who were approached, 620 completed the pen-and-paper questionnaire and returned it to a collection box in the clinic. Of these men, 523 self-identified as gay or bisexual or had been sexually active with a man in the past year and provided sufficient data to be included in the study.

HIV testing clinic sample

Many people seek an HIV test because they have been at risk of HIV infection. NHS clinics offer free voluntary counselling and testing for HIV and attract a broad cross section of people seeking an HIV test. From October 2002 to November 2003, men and women attending an NHS HIV testing clinic were invited to participate in the survey. Over the survey period, 1,889 people came to the testing clinic of whom 1,753 were eligible for the study. 1,640 were approached and 1,230 completed the pen-and-paper questionnaire and returned it to a collection box in the clinic. 334 men self-identified as gay or bisexual and provided sufficient data to be included in the study. A further 435 described themselves as heterosexual males and 450 as heterosexual females.

Recruitment was extended to a second NHS clinic which was specifically for gay men from June to November 2003. 211 men came for an appointment, of whom 209 were eligible. 198 were approached and questionnaires were completed by a further 156 men in this way, 70 of whom were seeking an HIV test and provided sufficient data for the study. This results in a combined HIV testing clinic sample of 404 men.

Questionnaire

Men in the gym, HIV positive clinic and HIV testing clinic samples completed a pen-and-paper survey whereas men in the Internet sample completed a web survey. Data were collected on a range of sociodemographic variables and on sexual behaviour, including sex with partners met through the Internet, as well as other confounding factors for HIV risk behaviour such as recreational drug use.

Table 2.1: Background characteristics of London men in quantitative survey samples

	Internet 2002		Internet 2003		Gym 2002		Gym 2003		HIV positive clinic		HIV testing clinic	
	n	%	n	%	n	%	n	%	n	%	n	%
Age (median; range)	33	18 – 70	32	18 – 75	35	17 – 79	36	18 – 75	38	23 – 70	32	17 – 73
Ethnicity (white)	1117	91.7	526	89.3	821	90.4	484	89.6	467	89.6	342	84.7
Employed	1007	82.7	477	80.2	771	85.3	450	83.2	324	62.8	317	91.4
Higher education	810	66.5	377	63.9	761	83.8	425	79.7	328	65.7	261	72.9
Gay (vs bisexual)	1084	89.0	524	88.1	869	95.1	523	96.3	489	93.5	354	87.6
HIV positive	142	11.7	67	11.3	138	15.1	88	16.2	523	100.0	15	3.7
Uses Internet to seek sex	1040	85.4	544	91.4	400	44.4	280	52.0	223	43.6	186	46.0
Completed clinic survey	-	-	26	4.7	-	-	16	3.0	-	-	-	-
Completed gym survey	24	2.0	16	2.9	-	-	-	-	25	5.0	14	3.9
Completed 2002 web survey	-	-	-	-	-	-	54	10.2	27	5.4	26	7.3
Response rate	** [1]		** [1]		50 – 60% [2]		50 – 60% [2]		66% [3]		70 – 75% [4]	
Base (n)	*1218*		*579*		*914*		*543*		*523*		*404*	

Note: [1] it is estimated that less than 1% of men using the websites over survey period responded; [2] estimated as a percentage of gay men using the gyms over survey period; [3] as a percentage of eligible men attending clinic over survey period; [4] as a percentage of eligible men attending clinic1 (70%) and clinic2 (75%) over survey period

Source: The Internet and HIV study: design and methods (Elford *et al.* 2004a)

The same core questions were included in all four questionnaires but each survey contained additional questions which were relevant to that group. Both the pen-and-paper and web surveys contained no identifying information. Identifiers such as IP addresses were removed from the web survey responses before they were stored in the database.

Sample characteristics

Table 2.1 summarises some of the background characteristics of the four London survey samples. It shows how the men in the different samples were mostly young, white and identified as gay. There were some differences between samples in employment and education, and the recruitment location is clearly reflected in differences in HIV status and use of the Internet for seeking sex. Few of the men had participated in another Internet and HIV survey.

Although Table 2.1 provides information on all four Internet and HIV samples, only data from the Internet and gym samples were used in the research for this thesis. The Internet sample was used because the purpose of this thesis is to explore the use of the Internet as a data collection tool. Chapter 4 examines issues of respondent motivation for participation in the 2003 web survey and Chapter 5 is an exploration of respondent drop out from this survey. Chapter 6 provides a detailed analysis of how the characteristics of the 2002 and 2003 Internet samples compare to men drawn from a national probability sample. Chapter 7 is an analysis of item nonresponse in the web and pen-and-paper surveys completed by the 2003 Internet and gym samples, respectively. The gym sample was used to provide comparative data from a pen-and-paper questionnaire because it includes men who are HIV positive, negative and untested in a similar way to the Internet sample, whereas the HIV positive clinic sample is limited HIV positive men and the HIV testing clinic sample is limited to those who have tested for HIV.

2.3 Internet and HIV study: qualitative methods

At the end of each of the four surveys described above, men were asked whether they would be willing to participate in an in-depth one-to-one interview with a qualitative

researcher. If they agreed, they provided contact details and the qualitative researcher contacted them to invite them to interview. In this way, 65 gay and bisexual London men were recruited from the Internet sample, 23 from the gym sample, 20 from the HIV positive clinic sample and 20 from the HIV testing clinic sample, a total of 128 men.

All men from the offline samples (gym and clinics) were interviewed in traditional face-to-face qualitative interviews. 30 men from the Internet sample were interviewed face-to-face and 35 were interviewed in an online chat. Men were not randomly assigned to interview mode. Online interviews were conducted in private chats using gaydar or gay.com which only the interviewer and respondent could enter. Chapter 8 compares the use of face-to-face and online interviews. It also provides further information on the use of these methods in both the Internet and HIV study and the methodological research undertaken for this thesis.

The Internet and HIV qualitative interviews were divided into two phases. 24 interviews were conducted in phase one which examined how men used the Internet for seeking sexual partners and 104 interviews were conducted in phase two which focused on Internet-related risk behaviour.

Sample characteristics

Table 2.2 (on the following page) is a summary of background characteristics of the four qualitative interview samples. All men were London-based. As with the survey samples, HIV status and use of the Internet for seeking sex vary according to where the men were recruited. Only men recruited from the 2003 Internet sample participated in the methodological analysis of face-to-face and online interviews that is described in Chapter 8. This is because men recruited for the gym and clinic samples were all interviewed face-to-face and it was therefore not possible to make the same comparison within these groups.

Table 2.2: Background characteristics of London men in qualitative interview samples

	Internet (online)	Internet (ftf)	Gym (ftf)	HIV treatment clinic (ftf)	HIV testing clinic (ftf)
	n	n	n	n	n
Age (median; range)	32 20–63	39 21-60	35 24–51	38 31–59	40 25–66
Employed	30	23	18	9	18
Higher education	20	20	17	15	15
HIV status					
HIV positive	6	11	1	20	-
HIV negative	20	15	18	-	20
Never tested	9	4	4	-	-
Sex seeking					
Rarely or never online	6	0	11	13	14
Online and offline	10	14	5	3	4
Mostly oneline	19	16	7	4	2
Base (n)	*35*	*30*	*23*	*20*	*20*

Note: (ftf) indicates face-to-face interview and (online) indicates online chat interview
Source: The Internet and HIV study: design and methods (Elford *et al.* 2004a)

2.4 Ethical framework

The Internet and HIV study operated within the ethical framework of informed consent. The research protocol was approved by the following committees: Royal Free Hospital and Medical School Local Research Ethics Committee, the East London and the City Research Ethics Committee, and City University London Research Ethics Committee. The following begins with a summary of the protocol for the quantitative arm before describing the ethical protocol for the qualitative interviews.

The Internet and gym samples were provided with a written explanation of the study and asked to tick a box if they agreed to participate before completing the questionnaire. No personal identifiers were included on these questionnaires.

The HIV positive clinic sample were provided with written information about the research and given an explanation of the study by a trained member of the research team in the clinic. If they agreed to participate, they were asked to sign a written consent form before completing the anonymous questionnaire. All data were treated in the strictest confidence. Consent forms and questionnaires were held separately and contained no identifiers linking them to one another. An identical procedure was followed with the HIV testing clinic sample.

When respondents volunteered for an in-depth qualitative interview, they were provided with written information about the research as a hard copy or email attachment. If they were still happy to participate, they were asked to complete a written consent form before the interview took place. Face-to-face interviewees could sign the consent form at the time of the interview and online interviewees could respond by email to indicate that they were willing to participate. In both cases, the respondent's contact details were not linked to the interview transcript and personal identifiers were removed from all transcripts.

The ethics approval for the Internet and HIV study covered the secondary analysis and collection of additional data that was undertaken for the methodological research in the quantitative arm. Because the methodological research in the qualitative arm included a further round of qualitative interviews with men who had taken part in the Internet and HIV study, additional ethical approval was sought and obtained from City University. The ethics protocol was based on that used by the Internet and HIV study and is described in detail in Chapter 8.

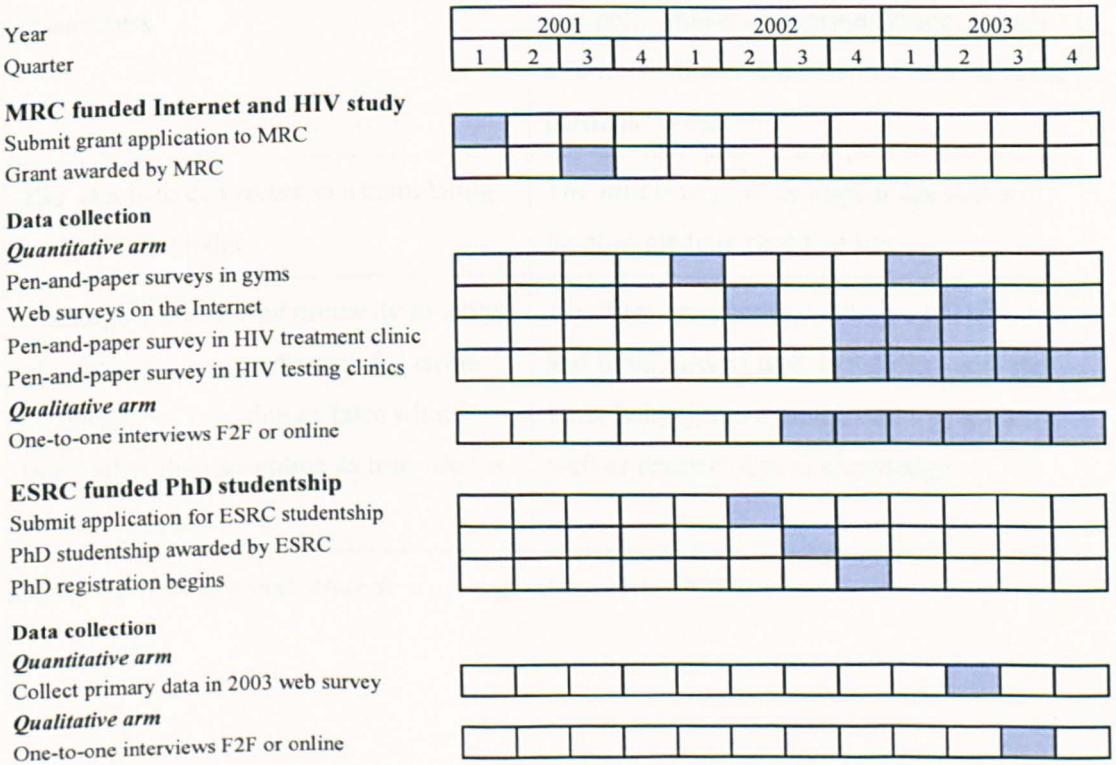
Thus, the framework of informed consent was central to both the main Internet and HIV study and the methodological research upon which this thesis is based. Although the subject of cyber-ethics has raised discussion among researchers in areas such as drawing the boundaries of online privacy, issues of online security and the identity of research participants, the approach taken here was to apply the framework of informed consent to Internet research in the same way that it has been applied in similar research using traditional methods.

2.5 The methodological research in context

The principal investigator on the main (MRC funded) Internet and HIV study was Jonathan Elford (City University). Other investigators were Graham Hart (MRC Social and Public Health Sciences Unit) and Lorraine Sherr (Royal Free and University College Medical School). Graham Bolding (City University) was the researcher responsible for the quantitative arm of the study and Mark Davis (City University) was responsible for the qualitative arm. Jonathan Elford was second supervisor of this doctoral thesis and the other members of the Internet and HIV research team were on the doctoral advisory board. All members of this board (including Dick Wiggins, first supervisor, and Roger Jowell, both at City University) agreed that the methodological research would benefit from adding questions to the 2003 Internet and HIV web survey and conducting follow-up qualitative interviews with men who had participated in the main Internet and HIV study in 2003.

Table 2.3 describes the chronology of the application for funding and data collection for the Internet and HIV study and the doctoral project.

Table 2.3: Chronology of key events for Internet and HIV study and doctoral project



3 Research strategy

3.1 Introduction

This methodological study began with a broad objective to *"to examine the issues surrounding the recent and rapid emergence of the Internet as a new research tool, in order to inform those wishing to undertake research via the Internet"*. It therefore represents an academic piece of research which, at the same time, has a highly practical objective. In order to situate the study, I will draw on Hammersley's (2000) typology of social research (Table 3.1), which makes three main distinctions between scientific and practical inquiry in terms of audience, aim and assessment of the findings.

Table 3.1: A typology of social research

Scientific inquiry	Practical inquiry
The immediate audience is fellow researchers	***The immediate audience*** is practitioners and policymakers of various kinds, as well as others who have a practical interest in the particular issue
The aim is to contribute to a cumulating body of knowledge	***The aim*** is to provide knowledge that will be of immediate practical use
Findings are assessed primarily in terms of validity; with a preference for erring on the side of rejecting as false what it true, rather than accepting as true what is false	***Findings are assessed*** in terms of relevance and timeliness as well as validity, with the latter being judged on the bases of lay as well as research-based knowledge

Source: Varieties of social research: a typology (Hammersley 2000: 227)

This study straddles the typology in a number of ways. Firstly, the target audience is fellow researchers who are also practitioners with an applied interest in the use of the Internet for the purpose of data collection. Secondly, the aim is both to contribute to an accumulation of knowledge and to provide practical solutions. On the third point, however, the study sits firmly on the side of scientific enquiry with critical evaluation of the findings being of paramount importance. Thus, the intention behind this methodological study is to communicate the findings of rigorous research to an audience of fellow researchers who will be able to make practical use of them.

It was originally conceived that the research would be undertaken mainly through secondary analysis of data from the Internet and HIV study in order to address questions about using the Internet as a data collection tool. An important aspect of the research design of this methodological study, therefore, was to capitalise on the large volume of data that had been collected both online and offline for the purpose of investigating high risk sexual behaviour among gay men. The methodological study provides added value by taking advantage of available data. In so doing, it does not conform to a text book approach to research whereby the research questions determine the design of the study. Although real life social research rarely conforms to such a text book approach (Bryman 1992), this methodological study is unusual in that the broad objective and available data were the starting point for the study to which specific research questions were then tailored.

Over the course of the study, further research questions were formulated that went beyond the limitations of the original Internet and HIV datasets. The issue of how Internet samples compare to probability samples was raised and access to data from Natsal 2000 was sought and granted, enabling additional secondary analysis of a sub-sample of MSM drawn from this probability sample (Chapter 6). As data collection for the Internet and HIV study continued into 2003, after the methodological study had started, it was possible to gain access to respondents whose personal experiences of Internet research could address the original broad objective. Primary research questions were thus formulated for which data were collected as part of the Internet and HIV study in 2003 (Chapters 4, 5 and 8).

Table 3.2: Research strategy: questions, methods and data

Chapter	Research question	Method	Data associated (primary data collection) or interrogated (secondary analysis)	Type of data
PART II – SAMPLING GAY AND BISEXUAL MEN				
4	Why do gay and bisexual men participate in web surveys of sexual behaviour?	Primary data collection	Internet and HIV web survey 2003	Quantitative
5	Among gay and bisexual men, who drops out of web surveys of sexual behaviour, at what point and why?	Secondary analysis Primary data collection	Internet and HIV web survey 2003[1]	Quantitative
6	How do self-selected Internet samples of gay and bisexual men differ from men who have sex with men drawn from a probability sample?	Secondary analysis	Internet and HIV web survey 2003 Internet and HIV web survey 2002 Natsal 2000	Quantitative
PART III – MODE EFFECTS IN STUDIES OF SEXUAL BEHAVIOUR				
7	What is the modal difference in item nonresponse when gay and bisexual men complete web and pen-and-paper surveys of sexual behaviour?	Secondary analysis	Internet and HIV web survey 2003 Internet and HIV pen-and-paper survey 2003 (gym sample)	Quantitative
8	What is the modal difference in interview experience when gay and bisexual men are interviewed face-to-face or using online synchronous chat?	Primary data collection	Internet and HIV qualitative interviews 2003	Qualitative

[1] Secondary analysis of Internet and HIV web survey 2002 included in Appendix 4

All the research questions stemmed from the original broad objective and were rooted in the relevant literature in a manner that will be described below and throughout the empirical chapters which follow. In this way, I have tried to avoid treating data from the Internet and HIV study as a *"data set in search of analysis"*, an approach to secondary analysis that is said to produce only trivial findings (Kiecolt & Nathan 1985: 14), and have aimed to provide a considered appraisal of systematic research questions that are rooted in the relevant literature.

Table 3.2 provides a summary of the research that was undertaken for this methodological study. It illustrates how the research questions form an integrated strategy to explore the data from the Internet and HIV study and Natsal 2000. The research questions are ordered according to the chapter within which they are examined, under the two major headings "Sampling gay and bisexual men" and "Mode effects in studies of sexual behaviour". The method used to investigate each question is described as involving primary data collection or secondary analysis. These methods are described in detail in the empirical chapters which follow. For those research questions that involved primary data collection, the Internet and HIV dataset with which the primary data were associated is named, while for those that required secondary analysis, the datasets interrogated are listed. The final column illustrates how most of this methodological study involves quantitative data but that the collection and analysis of qualitative data is also included.

The following describes the pragmatic approach to research and publications-as-theory strategy that were adopted in this study. It continues with a description of how the research design is situated in the wider literature and an overview of the advantages and disadvantages of secondary analysis. It concludes with a comment on the data analysis used in this study.

3.2 Adopting a pragmatic approach to research

The interplay between the broad objective of the study, the datasets and the research questions demonstrates how the scope of this methodological study was to a large extent shaped by the available data. It is important to describe how this relates to the

underlying methodology. In so doing, my aim is to justify what I believe to be the robust rationale for the design of the study. This explanation adopts Bryman's (1984) distinction between the use of "methodology" as an epistemological position and "method" as a technique of data collection.

The division that is most often applied to research in the social sciences is between quantitative and qualitative methods which are said to stem from opposing methodologies. As Table 3.2 illustrates, this study does not use quantitative and qualitative approaches to address the same research questions but uses both types of method to address different questions within the same study. Although the integration of quantitative and qualitative data is not therefore at issue here, the principles that lie behind the use of mixed methods are central to the approach that I have taken in this study.

Bryman (2004) summarises a simplified version of the fundamental differences between quantitative and qualitative approaches at the levels of the role of theory, epistemology and ontology (Table 3.3).

Table 3.3: Differences between quantitative and qualitative research strategies

	Quantitative	**Qualitative**
Principle orientation to role of theory in relation to research	Deductive: testing of theory	Inductive: generation of theory
Epistemological orientation	Natural science model, in particular positivism	Interpretivism
Ontological orientation	Objectivism	Constructivism

Source: Social Research Methods (Bryman 2004)

Thus, quantitative research is traditionally associated with a deductive approach to theory whereby research is used to test hypotheses, whereas qualitative research is associated with the inductive approach of generating theory that is grounded in the

research. A quantitative approach is said to derive from an epistemology based on the natural sciences model whereas qualitative approaches reject this model in favour of an interpretivist emphasis on the subjective meaning of social action (Bryman 2004). On the final point of ontological orientation, quantitative approaches are said to seek measurement of an objective, external reality which is incompatible with the qualitative approach whereby social phenomena are seen as socially constructed.

Although all research may be said to involve both deduction and induction in the interaction between ideas and data (Hammersley 1992), the described division between quantitative and qualitative approaches at the epistemological and ontological levels is difficult to reconcile. The problem with using mixed methods, therefore, is that these irreconcilable differences are assumed to apply to the methods with which they are associated. And yet, it is argued that we can divorce these opposing philosophical positions from the process of data collection because,

> *"research is a practical activity and cannot be governed in any strict way by methodological theory"* (Hammersley 1995: 19).

In support of this, Brannen (1992) describes real life social research whereby methodological theory appears to play little part in the actual choice of research methods. She found that she was drawn to the philosophical debate by the methodological literature and her own inclinations but that colleagues and research reports underlined the importance of pragmatic factors in the choice of research methods. Bryman (1984) has also argued that discussion of the philosophical basis of methodology is redundant to the extent that it does not solve the issue of which method is most suited to the research problem at hand. He finds that,

> *"there is no necessary 1:1 relationship between methodology and technique in the practice of social research"* (Bryman 1984: 89).

This study comprises an amalgamation of research methods. It is a combination of quantitative and qualitative approaches, online and offline methods and the analysis of primary and secondary data. The multi-method strategy that I have adopted is

consistent with what Tashakkori and Teddlie (1998: 12) have described as a pragmatic approach to research design whereby researchers do not concern themselves with the *metaphysical truth* but with *what works*. Such an approach emphasises the overriding importance of the research question (Tashakkori & Teddlie 1998). As previously described, access to data from the Internet and HIV study was a key factor in the design of this study. This takes the pragmatic approach beyond *what works* and into what can be made to work, acknowledging the inevitable trade offs associated with research in the real world. In this way, I tailored specific research questions and used methods that could be made to work with the resources that were available.

This approach recognises that different circumstances may generate other ways of investigating the research questions at hand and that the strategy adopted here represents what Silverman (2000: 825) has described as "*one way of 'slicing a cake'*". At the same time, it acknowledges that certain methods are better suited to certain ends. As will be explained in greater detail in subsequent chapters, when I wanted to find out why the thousands of respondents who completed the 2003 Internet and HIV web survey had done so, I added a limited number of closed questions to the end of the web survey. I tried to keep response burden to a minimum in order to encourage maximum participation. When I wanted to ask respondents about their experience of participating in online and offline qualitative interviews, I used online and offline qualitative interviews in order to encourage respondents to express their thoughts about the process.

Although I embrace the pragmatic approach to research, the influences of the epistemological and ontological perspectives are pervasive. Thus, one of the issues that arises from combining quantitative and qualitative methods is that of the author's voice. The tradition of the methods in this respect is different, with the researcher tending to situate herself within qualitative research through the use of the first person and to distance herself from quantitative research through adopting the third person. In writing up this research, it seemed fitting to adopt the third person in the quantitative chapters and yet, I fell naturally into the first person in the qualitative

chapter. As my aim throughout this thesis is to make my voice heard, it seems fitting that I should adopt the first person throughout, although the stylistic emphasis may vary from one chapter to another. In a personal communication, Bryman (May 2005) recommended the consistent use of one style of writing although he finds that the differences between the two styles are not so significant as they may first appear.

There is one further point arising from the quantitative-qualitative debate that I would like to raise before turning to the theoretical framework. Quantitative research has been criticised for ignoring the processes involved in data construction whereas qualitative research is said to focus as much on the process as the data (Holstein & Gubruim 1997). Since people communicate for a reason (Marsh 1982), all research communication results in data that are *produced* and not merely *collected* (Irvine *et al.* 1979: 3). One of the aims of this thesis is to investigate how data are moulded in order to contribute to the effectiveness of both quantitative and qualitative methods. As Marsh has argued, *"reactivity is a fact of life in social research"* (Marsh 1982: 56) that we cannot remove but must try to understand if we are to improve the way we work with our respondents.

3.3 Publications-as-theory strategy

My approach to this study most closely echoes that of the survey methodologists whose work forms the core literature here and who have traditionally tended to take a less theoretical and more empirical approach to the development of best practice for implementing surveys. The absence of a clear over-arching theoretical framework within which to ground all the aspects of this methodological study is consistent with an approach that Bryman (2004) has described as a *publications-as-theory strategy*. He makes a distinction between the grand theories that may be taught in sociology courses and what Merton (1967) has described as middle-range theories, which fall between grand theory and empirical findings, and which are often used to guide empirical inquiry. He does not, however, believe that social research need necessarily stem from such theory:

"It would be harsh, not to say inaccurate to brand as naive empiricism the

> *numerous studies in which the publications-as-theory strategy is employed, simply because their authors have not been preoccupied with theory. Such research is conditioned by and directed towards research questions that arise out of an interrogation of the literature ... The literature acts as a proxy for theory"* (Bryman 2004: 7).

The use of this strategy in the present study ensures that the research is thoroughly rooted in the current state of our knowledge on Internet research methods. It also acknowledges the diverse fields of research that are relevant to this particular study, which brings together literature from social research methods, sexual health and psychology. It is well suited to the exploratory secondary analysis of large quantities of data that was undertaken here, providing a clear framework within which to situate the research and guarding against the sort of data dredging or blind search for significant relationships that may be one of the potential pitfalls of doing secondary analysis with large datasets (Dale *et al.* 1988). Yin (1993) also recommends that exploratory studies such as this should set out a clear purpose which the research undertaken here has aimed to do, through use of the publications-as-theory strategy to turn the original broad objective of the study into a set of research questions.

The strategy makes it particularly important to ensure that the relevant literature has been thoroughly reviewed. The dynamic nature of the Internet makes it additionally paramount to keep abreast of recent, unpublished developments in the field of Internet research. In order to realise these two aims, the following methods of locating the relevant research were employed.

Two electronic databases were searched systematically using identical search terms that were designed to find references related to the range of topics covered by the thesis. The first database was the Social Sciences Citation Index, which contains references from journal articles in the social sciences from 1981 to the present day, and the second was PsycINFO, which contains references from journal articles, book chapters, books and dissertations in psychology and the psychological aspects of related disciplines from 1872 to the present day. A total of nearly two hundred searches were undertaken from which nearly five hundred references were located. Those references which proved to be of relevance to the study acted as a starting

point for identifying further relevant literature and experts in the field. The second method of locating recent, unpublished research was through attending and presenting work in progress at targeted conferences (see Appendix 1), attending other relevant conferences, seminars and workshops, and following up leads using the Internet.

3.4 The research design in a wider context

Research designs can be divided into cross-sectional, experimental, comparative, longitudinal and case study (Bryman 2004). The primary research undertaken here employed the cross-sectional design that is described below. The discussion of the design for the secondary analysis which follows illustrates the complications that arise in defining research involving post hoc comparisons and the extent to which they may be classified as experimental or comparative designs. The study did not utilise either of the remaining two designs: respondents were not tracked over time in the manner of a longitudinal design and the research fails to conform to Yin's (1993) definition of a case study. Thus, despite its reliance on detailed analysis of data from a single case, the Internet and HIV study, the thesis seeks to generalise to a broader population in a way that distances the findings from their immediate context and does not correspond to the case study approach.

Three of the research questions from the study were addressed through primary data collection alone (Chapters 4 and 8) or in combination with secondary analysis (Chapter 5), as illustrated in Table 3.2. Two of these questions used quantitative approaches and the third used a qualitative approach. A ***cross-sectional design*** was adopted in each case. Such designs are said to involve collection of data represented by two or more variables from a number of cases at a single point in time which are examined for patterns of association (Bryman 2004).

In order to address the questions raised in Chapters 4 and 5, two separate sets of survey questions about motivation for participation and reasons for dropping out of the survey were incorporated into the Internet and HIV 2003 web survey and a secondary analysis of survey drop out was undertaken. Although usually associated

with quantitative research, the cross-sectional design may also be applied to qualitative approaches involving unstructured or semi-structured interviews with a number of people (Bryman 2004) and was used in the design of the qualitative interviews undertaken for this methodological study (Chapter 8).

The design that is most often used in the examination of survey methods is experimentation. Typically, a survey is fielded and groups of respondents are randomly assigned to one of two or more experimental conditions which are manipulated while other conditions of survey participation remain constant. The randomisation process means that any effects that are found can be attributed to the experimental manipulation. When the researcher does not have full control over the conditions or the capacity to randomise this is known as a ***quasi-experimental design*** (Campbell & Stanley 1963), a number of which have been described (Campbell & Stanley 1963; Cook & Campbell 1979). Kish (1959), on the other hand, has argued that the term "experiment" should be reserved for cases where uncontrolled variables are randomised and, indeed, the use of the term "quasi" signifies a lack of the control that is the hallmark of an experiment (Marsh 1982).

The research questions in Chapters 6 and 7 of this thesis were addressed through secondary analysis of data from various samples of the Internet and HIV study and Natsal 2000, as outlined in Table 3.2. The aim was to assess how survey conditions impact on sample composition and item nonresponse. Whereas Campbell and Stanley (1963) have described this as a *static-group comparison*, which is a pre-experimental design, Cook and Campbell (1979) describe it as a *post test only design with non-equivalent groups*, which is a quasi-experimental design. These differences in categorisation illustrate the uncertain status of the quasi-experiment with regard to the design for this secondary analysis where comparisons were made between one group and a single non-equivalent group that was not subjected to the same experience.

In both cases, the design is graphically represented by the following, where X is the survey condition, O is the observation and the dotted line indicates the lack of

equivalence between groups (Campbell & Stanley 1963):

$$\begin{array}{ccc} X & & O_1 \\ - & - & - \\ & & O_2 \end{array}$$

The above design implies the presence or absence of X, whereas the observations undertaken for this study involved comparison of two different survey conditions. Although this is consistent with the description given by Shadish *et al.* (2002), whereby the comparison group receives an alternative intervention or no intervention at all, the lack of control over the survey conditions in addition to lack of respondent randomisation in the present study, may make it more accurate to abandon any pretensions to experimentation in favour what Bryman (2004) has called a ***comparative design***. This design is most often used in the context of cross-national research but may be employed in other comparisons where the same method, or in this case the same analysis, is applied to two contrasting cases for the purposes of comparison. It may be expressed by the following comparison of non-equivalent groups where one experiences condition X and the other experiences condition Y:

$$\begin{array}{ccc} X & & O_1 \\ - & - & - \\ Y & & O_2 \end{array}$$

Whether subsumed under a quasi-experimental or comparative design, the important point to note is that the design lacks the control to make causal inferences which is the unique strength of an experimental design (Shadish *et al.* 2002). Indeed, neither this nor the cross-sectional design described earlier can be used to confirm causal inference and may thus be described as correlational. As Campbell and Stanley (1963) have argued, however, although correlation cannot establish causation, it is the outcome of the causal laws that produce its effects. Thus, it is possible that differences found between O_1 and O_2 may be attributable to the conditions X and Y but because it is not proven, plausible alternatives should be given due attention (Shadish *et al.* 2002). Even experiments are more likely to provide causal

description than causal explanation (Shadish *et al.* 2002), however, and the fundamental problem of such description is that,

> *"the assumed meaning of the relationship, how it works mechanically or symbolically, is not in the data but laid on top of it"* (Marsh 1982: 102).

Given the above obstacles to the attribution of cause and meaning in the present study, the interpretation of the findings is necessarily speculative.

3.5 Secondary analysis

Because secondary analysis forms a significant part of the research strategy, the following section provides an overview of the advantages and disadvantages of its use in the present study. A detailed description of how it was used will be given in the following empirical chapters which will also detail the other methods used.

Secondary analysis may be carried out on data that have been previously analysed for a different purpose, pre-existing data that have not before been analysed or data that have been amalgamated into a new dataset (Dale *et al.* 1988). It is used to address research questions that were not anticipated at the time of data collection and tends to be carried out by someone who was not involved in collecting the data (Bryman 2004). In this study, all secondary analysis was carried out by myself on data that were collected by other researchers and had previously been analysed for the purpose of examining sexual health and behaviour.

Before embarking on secondary analysis, data require special scrutiny to determine the extent to which they fit the purpose of the analysis (Dale *et al.* 1988; Stewart & Kamins 1993). Stewart and Kamins (1993) have identified six questions which should be asked for this purpose. They essentially cover issues relating to the quality of the data and highlight the importance of transparency in the data collection process for the assessment of the data:

1. What was the purpose of the study?

2. Who was responsible for collecting the information?
3. What information was actually collected?
4. When was the information collected?
5. How was the information obtained?
6. How consistent is the information obtained from one source with information available from other sources?

Questions one to five were addressed in the earlier description of the Internet and HIV study and the answer to question six is the subject of Chapter 6. The above questions will be similarly addressed for data from Natsal 2000 in Chapter 6. The issues are thoroughly covered in the published articles on the findings from the Internet and HIV study and Natsal 2000, and in the detailed documentation which is available at: http://www.biomedcentral.com/1471-2458/4/39 (Internet and HIV study) and http://qb.soc.surrey.ac.uk/surveys/nssal/nssalintro.htm (Natsal).

The Economic and Social Research Council (ESRC) has been concerned that large-scale datasets are under-utilised and, since 2001, has set aside funding to support twenty PhD students annually under the Secondary Analysis Studentship Scheme and has recently launched a Secondary Data Analysis Initiative. The research council recognises the potential for secondary analysis but also acknowledges the challenges which require researchers,

> *"to understand the conditions under which the data were collected so that they may be correctly interpreted, to devise ways of making imperfect data substitute for primary data, and to be imaginative and skilled in applying and developing methods that take account of these factors but still address the questions at hand"* (ESRC 2005).

Although all data have their imperfections, the major limitation of secondary analysis is that surveys rarely contain the precise information that is sought and secondary analysts must be mindful of the extent to which the questions measure their concepts of interest (Hyman 1972). When two or more datasets are included in the analysis, issues of question and sample comparability also arise (Kiecolt & Nathan 1985). The use of several datasets and samples in this study meant that comparability was

certainly an issue although the driving force behind the analysis was to locate equivalent measures of equivalent concepts rather than to find measures for specific, pre-determined concepts.

The equivalence of samples is the subject of Chapter 6 and the issue of question comparability will be described and addressed in the methods section of the relevant chapters. An example of secondary analysis undertaken by Smith (1979) will, however, serve to demonstrate how small differences in question wording, format and context may influence such comparability. Smith used data from a number of surveys measuring self-assessed happiness in order to examine trends in psychological well-being in the US since the Second World War. When respondents were given a choice between "very", "fairly" and "not very happy" they were more likely to choose "very happy" than respondents who could choose between "very", "pretty" and "not too happy". The addition of "completely" and "not at all happy" to the extremes of the latter scale had little effect on the unhappy end of the continuum, but it shifted response upwards at the happy end from "pretty" to "very happy" and from "very" to "completely happy". Furthermore, when the happiness question was placed after a question about marital happiness it appeared to raise reported happiness for married respondents. Thus, the use of different words, response formats and contexts were all found to influence response.

The lack of question equivalence is certainly an important proviso in doing secondary analysis but there are also a number of advantages associated with it. Using existing survey data saves money, time and personnel (Hyman 1972). As a student on a limited research budget and a novice researcher, I would not have had the resources to collect data of similar quality and quantity and have furthermore benefited from the expertise of those responsible for its collection (Dale *et al.* 1988). The fact that secondary analysis can be undertaken independently (Kiecolt & Nathan 1985) is also ideal for PhD research. Such savings are furthermore to the public good, where scarce resources may be directed towards essential new research (Hyman 1972).

Another advantage is that secondary analysis requires no further intrusion into people's lives (Hyman 1972). Although this may cloud the issue of informed consent and open up the possibility for data to be used in a way that is detrimental to the respondent (Dale *et al.* 1988), it is also in keeping with the ethical position that only required or essential research should be undertaken (Hyman 1972) and may signify greater respect for the respondent's valuable contribution through its fuller use. Secondary analysis may also benefit the wider research community and its beneficiaries through not adding to the over-surveying effect that has been implicated in the falling response rates over the past few decades. For example, some four and a half thousand UK men completed the Internet and HIV web survey in 2002, whereas only two and a half thousand did so in 2003, indicating that it may be unwise to over-burden this valuable resource and well-advised to make thorough use of the data that they provide.

The use of secondary analysis is complementary to the pragmatic approach to research described earlier. Both are concerned with the practical issue of making the most of what is available but whereas the pragmatic approach may be said to open up the range of methods available, secondary analysis may be limited by what can be achieved with the available data. As Hyman (1972) points out, however, it is the ability to realise the potential of the data that is important rather than the limitations with which the secondary analyst starts. In this way, the aim of the research strategy adopted here was to realise the potential of working with the Internet and HIV study datasets and researchers.

3.6 Data analysis

A number of different data analysis techniques were selected for their capacity to address the research questions in the empirical chapters which follow. The methods section of each chapter contains a detailed description of the techniques used in order to provide a framework for the findings which follow.

PART II

SAMPLING GAY AND BISEXUAL MEN

4 Web survey motivation for participation

Synopsis

Primary survey data were analysed to examine respondents' motivation for participation in the Internet and HIV 2003 web survey. Contributing to research and the survey topic were the most important motivating factors, regardless of individual characteristics, and participation appeared to be driven by respondent altruism.

4.1 Introduction

This exploration of the use of the Internet as a data collection tool begins with an examination of what motivates men to participate in such surveys in the first place. It focuses on motivation for participation in the 2003 Internet and HIV web survey and aims to address the question:

"Why do gay and bisexual men participate in web surveys of sexual behaviour?"

The following invitation (Figure 4.1) to participate in the Internet and HIV web survey appeared on pop-ups and banners in gaydar and gay.com, from May to June 2003.

Figure 4.1: Internet and HIV 2003 web survey banner

Over the five-week period that this invitation was extended, two and a half thousand men living in the UK completed the survey. Although click-through rates from banner advertisements are generally less than 2% (Tuten *et al.* 1999) and the

response rate here was estimated to be no greater, the large number of respondents recruited via this advert raises the question as to what motivated these men to respond in the first place.

Goyder (1987) highlights the multidimensionality of the response process and describes how survey participation is influenced by variables ranging from respondents' deeply held views to their immediate circumstances at the time of the survey request. Similarly, Groves *et al.* (1992) argue that an integrated approach to survey participation should be taken which draws on a range of influences. Those relating to self-completion questionnaires are described as societal level factors, attributes of the survey design, characteristics of the sample person and psychological factors. The type of circumstantial factors to which Goyder (1987) refers are also likely to have an important influence but one that is difficult to measure and is rarely explored in the literature. Their impact on survey drop out will be considered in Chapter 5.

Although this thesis will argue that the Internet has great potential for studying sexual behaviour among gay and bisexual men, Catania *et al.* (1993) have found that the issue of motivation for participation in studies of sexual behaviour is poorly understood. The present chapter aims to contribute to our understanding of this area in order to improve survey design and aid interpretation of our results (Sandfort 1997). It aims to empirically examine the following type of assertion that has been made about the motivation for participation in web surveys since their early days:

> *"Self-selected respondents probably participated because the survey was on-line, easy to access, and easy to answer. Also, the survey was relevant to a community in which they were involved and to topics about which they had knowledge and opinions"* (Walsh *et al.* 1992: 244).

Groves and Couper (1998) have argued that sociodemographic factors may not directly affect the decision to participate but may produce certain psychological predispositions that influence participation. Similarly, Goyder (1987) believes that attitudes towards survey participation are largely independent of social role. The

focus of this chapter, therefore, is on the motivational forces that influence participation. It brings together the current understanding of what motivates respondents to participate in surveys in general, in surveys of sexual behaviour and in Internet surveys, to undertake a unique investigation of what motivates gay and bisexual men to participate in web surveys of sexual behaviour. An understanding of this issue may aid survey design, given that,

> *"It is generally accepted that surveys would be better planned and implemented if we, the researchers, knew why people respond to requests for survey participation as they do"* (Evangelista *et al.* 1999: 241).

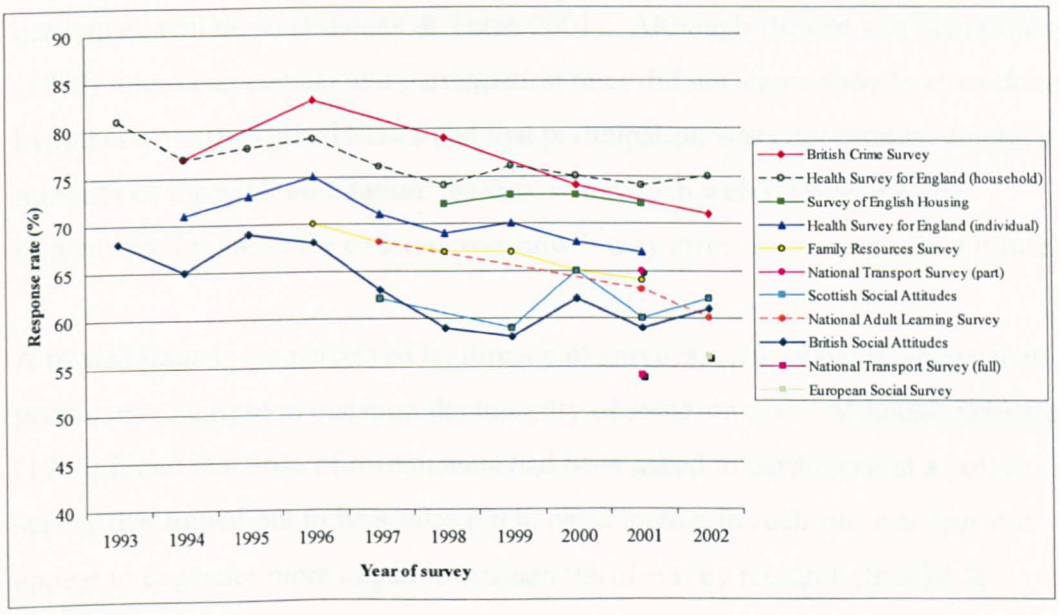
This introductory section begins with a brief outline of the impact of societal level factors on survey participation in order to set the broader context. It goes on to consider theoretical frameworks that have been suggested for participation motivation at the individual level, including Cialdini's model of compliance (Cialdini 2001) that was used to frame the present study. The section continues with a brief summary of the literature on how incentives have shaped participation. The main part of the section frames findings from empirical studies of the psychological, behavioural and survey design influences on survey participation in the context of Cialdini's model of compliance.

4.1.1 Factors that may influence participation

4.1.1.1 Societal Level Factors

One of the main concerns for the survey methodologist today is that of falling response rates. A recent study in the United States reported that response rates to the University of Michigan's Survey of Consumer Attitudes had fallen from 72% in 1979, to 60% in 1996 and 48% in 2003 and highlighted the continuing and increasing decline over the past few years (Curtin *et al.* 2005). Figure 4.2 (on the following page) illustrates the less dramatic but similarly worrying trend for a number of national UK-based interview-conducted surveys over the past decade.

Figure 4.2: Response rates for UK-based national surveys 1993-2002



Source: National Centre for Social Research

Goyder (1987) noted that the fall in response rates to US interview-conducted surveys in the 1960s and 1970s was not matched by a similar trend in postal questionnaires and attributed increasing non-response in part to the invasion of privacy that the individual may feel at the approach of the interviewer. In this respect, the increase in non-response may be symptomatic of a wider societal level change in attitudes towards personal privacy, which may now extend to the virtual world. Thus, requests for survey participation via unwanted pop-ups or unsolicited emails are the online equivalent to real world approaches that may generate similar frustration among potential respondents (Mehta & Sividas 1995).

Another societal level explanation of falling response rates is the over-surveying effect (Groves *et al.* 1992). The number of surveys that people may be faced with nowadays is believed to result in a lack of willingness to participate. Schliefer's (1986) examination of attitudes toward and participation in survey research found that nearly one fifth of respondents had been interviewed at least four times in the past year and Goyder (1986) found that the attitude towards research becomes more negative with increased number of requests to participate. In the face of the over-

surveying effect, it is argued that only those respondents who have a propensity to participate will do so (Bosnjak & Tuten 2001). Although Bickart and Schmittlein (1999) found that contact and participation rates did not support the "survey deluge hypothesis" in the United States and that participation was concentrated among a minority of the adult population, the ease with which web surveys may be implemented is cause for concern over how it may affect response rates in future.

A related issue is the perceived legitimacy of surveys. The evidence shows that people may be right to question the integrity of some surveys. Although Schleifer (1986) found that 36% of respondents had been asked to participate in a poll or survey that turned out to be a sales pitch, participating in such surveys does not appear to engender more negative evaluations of survey research (Stocké & Langfeldt 2004). It is to be hoped that such approaches will continue to be dissociated from legitimate surveys and not contribute to the type of blanket cynicism towards surveying which Joan Bakewell (2003) encapsulates in her article, "*Nine out of ten surveys are rubbish*".

Respondents may also be influenced by the environment in which they live. People who live in large cities are generally less likely to respond than those living in small towns, who in turn tend to be less cooperative than those living in rural areas. An examination of response rates in six surveys confirmed that cooperation was lower in urban, densely populated, high crime areas (Couper & Groves 1996). Although household level variables accounted for half of this effect, the data indicated that the social environment was independently associated with likelihood of response.

It is likely that changing societal attitudes towards sex may impact upon participation in a survey of sexual behaviour. The increased prevalence in many reported sexual behaviours between the 1990 and 2000 fielding of Natsal, were attributed to an actual change in these behaviours and more tolerant social attitudes, as well as improved survey methodology (Johnson *et al.* 2001b). Such societal level changes in attitude and behaviour might be expected to result in increased response rates for surveys of sexual behaviour, given that respondents to these surveys are more likely

to have liberal sexual attitudes and behaviours (Dunne *et al.* 1997). Changing attitudes towards homosexuality might also influence response to surveys about sex between men. Such changes are illustrated by the Natsal finding that 54% of women and 67% of men considered sex between men to be wrong in 1990, compared to 28% of women and 49% of men in 2000 (Erens *et al.* 2003).

Also at the societal level, attitudes towards the Internet and its various uses are likely to change as it becomes less of a novelty and more an integral part of people's lives. Although increased use of the Internet will undoubtedly reduce the problem of coverage, it remains to be seen how this will impact people's attitudes towards use of the Internet as a data collection tool.

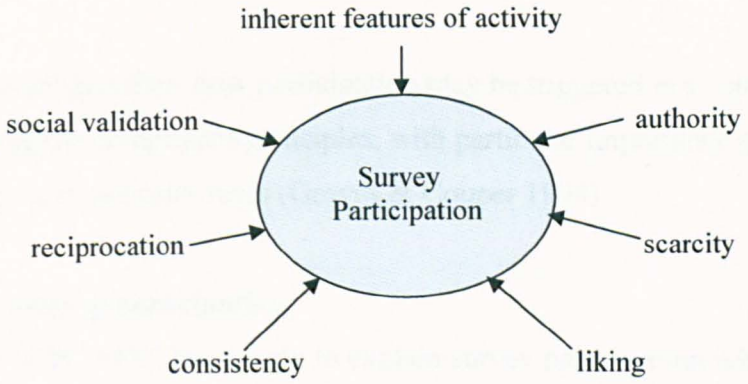
4.1.1.2 Individual Level Factors

Although an integrated approach to survey participation (Groves *et al.* 1992) highlights the importance of a range of influences, this chapter concentrates on those individual factors of relevance to the 2003 Internet and HIV web survey. The following outlines the motivational framework that was adopted in the present study and explores its relationship with other models of response and the empirical evidence.

Cialdini's model of compliance

Groves *et al.* (1992) propose that Cialdini's model of compliance (Cialdini 2001) may be usefully adopted to explain survey participation at the individual level. The model (Figure 4.3 on the following page) was originally developed as description of the principles governing the ways in which "compliance professionals", such as salespeople, fund raisers or advertisers, achieve their goals (Cialdini *et al.* 1999). It can also be used to describe how the inherent features of an activity, plus the six compliance principles of reciprocation, consistency, social validation, authority, scarcity and liking may determine whether an individual participates in a survey.

Figure 4.3: Cialdini's model of compliance



The *inherent features* that may attract or deter survey participants are considered here to relate to the topic that it covers and the act of questionnaire completion. The principle of *reciprocation* describes the sense of obligation that may be engendered when the respondent feels in some way indebted to the researcher. The principle of *consistency* may come into play if participation is consistent with the respondents' attitudes, beliefs or earlier actions. Where the respondent believes that people with whom he or she identifies would also participate, the principle of *social validation* may apply. Thus, social validation is invoked through reference to others whereas consistency is invoked with reference to the self. The principle of *authority* describes how the respondent is more likely to comply if the survey request comes from what he or she deems to be a proper authority. If the survey appears to represent a rare opportunity to participate, the principle of *scarcity* may be invoked. Factors that increase the respondents' *liking* for the survey interviewer may increase their likelihood of participating. Adapting this principle to self-completion surveys, Dillman (2000) has argued that people may be more likely to complete an attractive self-administered questionnaire than an unattractive one.

Groves *et al.* (1992) favour the heuristic approach described by this model, arguing that respondents do not devote time or cognitive energy to the decision to participate in a survey and that,

"few householders have strongly preformed decisions about survey requests.

Rather, these decisions are made largely at the time of the request for participation" (Groves & Couper 1998: 32).

The approach describes how participation may be triggered at a subconscious level by invoking the compliance principles, with particular importance placed on the reciprocity and authority rules (Groves & Couper 1998).

Other theories of participation

Dillman (1978; 2000) has sought to explain survey participation with reference to social exchange theory. According to social exchange, increasing respondent reward and decreasing respondent burden through engendering respondent trust is more likely to result in participation. Trust may be engendered through the sponsorship of credible organisations such as universities (Dillman 1978). The reward is not limited to the respondent alone but may extend to those with whom he or she identifies (Dillman *et al.* 1996). Groves *et al.* (1992) argue that Cialdini's reciprocation rule should be seen in the context of social exchange.

Sharp and Frankel (1983) found that perceived burden was associated with longer questionnaires, although those who took part in survey interviews did not generally consider them to be burdensome. The salience of the survey topic is also key to encouraging response, with a positively salient topic invoking the principles of social exchange. Dillman (2000) lists the practical ways in which the researcher may improve survey design in order to realise the aims of increasing respondent reward, decreasing perceived burden and creating trust.

Social exchange should not be confused with economic exchange, as the following investigations of the use of incentives will describe, but when the task of completing the survey is particularly onerous, respondents may require compensation for their time and effort (Tourangeau 2004). Thus, Knowledge Networks' provision of WebTV equipment in exchange for regular survey participation may be seen in terms of economic exchange (Chang & Krosnick 2003). It is suggested that other online panels may require compensation for the cost of their dial up connection to the Internet (Göritz 2004).

The leverage-salience theory of participation has sought to explain the relative importance of different appeals for participation to the individual (Groves *et al.* 2000). The theory proposes that people will vary in the importance that they attach to different aspects of the survey request (leverage) and this will impact on the likelihood of their response to the extent that such aspects are highlighted in the request (salience). The positive effect of an incentive was thereby found to diminish when respondents were more highly motivated by the topic of the survey (Groves *et al.* 2000), in accordance with the theoretical position that incentives will have the greatest impact on those who are the least likely to participate from other motivations (Tourangeau 2004).

The leverage-salience model is consistent with the elaboration likelihood model (Petty & Cacioppo 1984) which proposes that respondents who are engaged in a task take a central route based on the quality of the argument and those who are not involved take a peripheral route based on the quantity of the arguments. Thus, different appeals may invoke different levels of information processing, with an interest in the survey topic resulting in more extensive processing of the survey request (Groves *et al.* 2004). Cialdini's model highlights the less deliberative aspects of the decision-making process (Groves & Couper 1998), whereas the leverage-salience model suggests a mechanism whereby certain survey appeals may carry more or less weight for particular individuals. Although the two theories do not appear to be incompatible, the role of each within a given situation is yet to be tested.

Use of incentives

The use of monetary incentives and follow-up contacts are generally found to be the most effective ways of increasing response rate (Helgeson *et al.* 2002). Although follow-up does not apply to convenience samples and, in common with many British surveys, no incentives were offered in the 2003 Internet and HIV web survey, there is a large body of research into the impact of incentives on the likelihood of response which merits brief discussion here. It should be noted that this research has been mainly conducted in the USA and may have less relevance in the UK which does not have the same tradition of offering financial incentives. It has furthermore been

found that receipt of financial incentives for survey participation did not reduce respondents' likelihood of subsequent participation where no such incentives were offered (Singer *et al.* 1998).

Although no one-size-fits-all when it comes to the use of incentives (Albaum *et al.* 1998), certain strategies have generally been found to increase response rates. Prepaid incentives are consistently shown to be more effective than promised incentives and cash is more successful than nonmonetary inducements (Church 1993; Edwards *et al.* 2002). Thus, inclusion of $1 and $5 cash incentives resulted in response rates of 64% and 71%, whereas the promise of a $50 resulted in a lower response rate of 57%, that was barely higher than the no incentive rate of 52% (James & Bolstein 1992). Studies have also shown the diminishing returns of increasing the monetary value of incentives, with a recent study finding that increasing the incentive to ten Canadian dollars showed no improvement over five (Goyder & Brown 2004).

Dillman (2000) thereby maintains that the best alternative is to include a prepaid token financial incentive of a dollar or two. The effectiveness of such small, prepaid incentives can be explained in terms of social exchange, with the token prepayment invoking the norm of reciprocity and reinforcing the respondent's involvement in the research relationship (Collins *et al.* 2000).

One of the challenges for Internet surveys is to find electronic equivalents to the incentives that have been found to be effective in traditional surveying (Couper 2001). Although incentives may be successfully sent via other communication modes (Kypri & Gallagher 2003), it is not easy to replicate the inclusion of a token financial incentive in the online environment. Thus, an email survey with attached gift certificate secured a response rate of 32%, compared to postal survey response rates of 40% with attached gift certificate and 57% with enclosed cash incentive (Birnholtz *et al.* 2004).

Although a meta-analysis of panel and one-off web surveys found that incentives encouraged response (Göritz 2004, cited in Göritz 2004), there is no consistent evidence on the most effective method. They have furthermore been found to have an occasional impact on substantive analyses (O'Neil 2002) and should be used with care. Bosnjak and Tuten (2003) achieved a response rate of 65% in a web survey offering entry to a prize draw, with lower response rates achieved for respondents offered $2 through PayPal, the online money transfer service, in both the promised and prepaid conditions (58% and 56%, respectively). The authors highlight the possible burden and suspicions that respondents may feel about collecting money via PayPal, which may have contributed to the lower response rates associated with the use of incentives in a meta-analysis of electronic surveys (Cook *et al.* 2000). The market research company Virtual Surveys reported that the use of incentives made little difference on response rates to their pop-up surveys (Comley 2000).

The appeal of prepayment was, however, found to hold with response rates of 79% achieved when respondents were offered a $15 online gift certificate before participation compared to 69% when the certificate was offered on completion (Downes-Le Guin *et al.* 2002). Increasing the prepayment to $25 did not increase the response rate. As previously discussed, much of the research into incentives has been conducted in the USA and this chapter seeks to examine what encourages respondents to participate online in the absence of incentives, in accordance with the UK tradition.

4.1.2 Empirical findings on respondent compliance

4.1.2.1 Survey topic

Goyder (1986: 39) refers to "*the elasticity of survey response according to topic*" and the survey topic is generally acknowledged to be one of the most important influences on response behaviour (Goyder 1987; Groves *et al.* 1992; Vehovar *et al.* 2002).

Level of topic interest was found to have a key influence on response across a range of studies (Edwards *et al.* 2002), with a meta analysis of postal questionnaires

finding that response rates increased from 42% to 77% when the topic was judged to be more salient to the respondents (Heberlein & Baumgartner 1978). High interest surveys have also been shown to achieve higher response rates than low interest surveys within the same population (Martin 1994; Van Kenhove *et al.* 2002) and when respondents were asked about their willingness to participate in surveys covering seven different topics, they were least likely to refuse a survey about their telephone company (14%) and most likely to refuse a survey about voting intentions (51%) (McDaniel *et al.* 1987). Although certain survey topics are therefore more likely to secure higher response rates, people who respond to these surveys are also likely to have a greater interest in the topic. Thus, respondents and nonrespondents in a low interest survey on fast food had similarly low levels of topic involvement, whereas respondents in the high interest survey on clothing had a greater interest in clothing than their nonresponding counterparts (Van Kenhove *et al.* 2002).

The importance of the survey topic has also been found in studies using electronic surveys, with a meta analysis of such surveys finding that response rates were slightly higher with increased topic salience (Cook *et al.* 2000). Bosnjak and Batanic (2002) adopted Porst and von Briel's (1995 cited in Bosnjak & Batanic 2002) model in their exploration of willingness to participate in a scientific email survey. This describes three sets of reasons for survey participation – altruistic reasons, survey-related reasons and personal reasons. They found that curiosity about the subject of the survey (survey-related) was the most important motivating factor in responding to the survey, followed by contribution to knowledge (altruistic), self-knowledge (personal) and material incentives (which they added to the original model). In relation to Cialdini's compliance model, curiosity about a survey topic is considered to be an inherent feature of the survey and the evidence indicates that the sexual content of the Internet and HIV web survey is likely to have an important impact on who participates. The desire to contribute to knowledge, develop one's own self-knowledge and receive material gain can all be understood as elements of the reciprocation principle.

4.1.2.2 Reciprocation

Bosnjak and Batanic's study illustrates the complex nature of the reciprocation principle and the various ways in which it is invoked to reward respondents for survey participation. A useful way of conceptualising the difference between such rewards is to distinguish between motivation which is intrinsic to the act of survey participation and that which is extrinsic. In a banner-advertised web survey, it was found that an intrinsic appeal to contribute to an important study generated significantly more click throughs than and extrinsic appeal to win valuable prizes (Tuten *et al.* 1999).

Evangelista *et al.* (1999) used four theoretical frameworks to devise their list of motivational items for participation in a market research survey and their findings highlight the importance of the survey topic and Cialdini's reciprocation principle. The first framework of cognitive dissonance (Festinger 1957) states that behaviour that is inconsistent with the individual's self-perception as a helpful person induces an uncomfortable psychological state, called dissonance, which individuals aim to avoid. Similarly, the theory of self-perception (Bem 1967) predicts that an individual's behaviour is modified in order to be consistent with perceptions of the self. The third framework of commitment / involvement (Becker 1960) describes how the individual's involvement will determine their degree of commitment to a behaviour (Albaum *et al.* 1998). Social exchange is the final framework, as previously described. The most important factors were the topic of the survey which was used to operationalise commitment / involvement and a monetary incentive derived from social exchange.

Although appeals to altruism and social utility do not directly invoke a norm of reciprocity (Dillman *et al.* 1996), respondents who comply with such appeals may derive their reward from the personal satisfaction associated with conforming to the norm of social responsibility, especially in the context of socially relevant surveys (Warriner *et al.* 1996). The appeal of general societal contributions, however, may be less effective than contributing to a sector of society in which the respondent has a personal interest. Thus, response rates were not improved by offering charitable

donations for participation (Warriner *et al.* 1996) or highlighting the social utility of responding to the US census (Dillman *et al.* 1996). It is also possible that explicit appeals have little benefit over those that are implicit, with a systematic review of the literature finding that the inclusion of appeals stressing benefits to the respondent, the sponsor or society had little effect (Edwards *et al.* 2002). Perceived usefulness has been found to be an important factor in respondents' attitudes towards web surveys (Huang & Liaw 2005) and gay men might participate in the Internet and HIV web survey because of an implicit understanding that their contribution is reciprocated by practical research which aims to improve conditions for the gay community. In this way, Hispanic MSM stressed the importance of curiosity and altruism as their motivation for participation in a study of high risk sexual behaviour among their peers (Fernandez *et al.* 2004).

4.1.2.3 Social validation

The respondents' affiliation with the gay community may also invoke the principle of social validation whereby respondents comply to the extent that people with whom they identify might also do so. One method of suggesting this principle to respondents is through linking the survey request to websites or listservs that are aimed at the group of interest. A study of HIV risk behaviour found that the most effective method of recruiting respondents was to publicise their survey on gay and lesbian-orientated websites and listservs that were specifically aimed at their target groups (Epstein & Klinkenberg 2002). Market research web surveys elicited higher response rates where people had a strong relationship with the brand, such as a pop group website or the Open University (Runham 1999 cited in Comley 2000). Respondents may thus be more likely to participate in the Internet and HIV web survey because it is reached via gaydar and gay.com.

4.1.2.4 Liking

The research tends to suggest that the appearance of the survey is not likely to have a strong impact on the likelihood of response, indicating that the principle of liking may be less relevant to self-completion surveys than interviewer surveys. There was no substantive difference between initial compliance rates for respondents asked to

complete web surveys that were "plain" compared to "fancy" (Dillman *et al.* 1998) and an investigation of rates of click-through to banner adverts found that design was of secondary importance to the nature of the audience and their interest in the product (Briggs & Hollis 1997). For a postal survey, a positive attitude towards research and time spent dealing with post increased attention paid to the survey whereas manipulation of colour and personalisation features did not (Helgeson *et al.* 2002). Although respondents were more likely to complete a web survey when the email request for participation contained a photograph of a researcher rated as reasonably attractive than when the email contained no photograph (Guéguen & Jacob 2002), the use of such online images is probably unwise until their impact on response has been thoroughly examined (Couper *et al.* 2004; Witte *et al.* 2004). These findings suggest that it is unlikely that the look of the Internet and HIV web survey will have a strong influence on respondents.

4.1.2.5 Consistency

The above studies have shown how respondents behave in a way that is consistent with their beliefs and personal identification, with the superficial elements of survey design having less of an impact. According to the principle of consistency, they may also respond because this is consistent with previous response behaviour. Nederhof (1986) found that respondents who had more research experience tended to respond with less hesitation, whereas those with less experience often needed to be persuaded by follow-ups. Such behaviour may be related to a positive attitude towards both the value and enjoyment of survey completion, which were found to be associated with a general willingness to participate in future research (Rogelberg *et al.* 2001), and a positive attitude towards research, which has been found to encourage response (Helgeson *et al.* 2002).

These findings indicate that respondents to the Internet and HIV web survey are likely to have participated in other surveys. Similarly, participants in a sex survey were more likely to have been involved in at least one survey in the past 6 months (Turner 1999) and HIV Futures III, a survey of Australian HIV positive respondents, found that nearly two thirds of respondents had participated in at least one of the

research projects listed and nearly half had participated in earlier rounds of the HIV Futures project (Grierson *et al.* 2002).

The above study shows how people may respond to surveys in a way that is consistent with their previous behaviour, although such surveys may also invoke the principle of social validation whereby participation is consistent with one's peers' behaviour. Consistency thereby taps into individualism and social validation into collectivism. Thus, consistency was found to be a more effective appeal among individualists and social validation among collectivists and, while social validation and consistency were both motivating forces in samples of US and Polish students, there was a stronger emphasis on consistency among the US sample and on social validation among the Poles (Cialdini *et al.* 1999).

Previous participation in Internet surveys has also been found to be one of the most important determinants of participation in a subsequent Internet survey (Batagelj & Vehovar 1998). Such participation may be consistent with an individual's more extensive use of the Internet in general. University students' response to a web survey was significantly related to their Internet use and familiarity, controlling for gender, age and year in school (Kwak & Radler 2002). 86% of subscribers with higher than median network usage responded to an online survey request compared to 66% of below median users (Walsh *et al.* 1992) and logfiles tracking Internet usage showed that respondents to an email survey were strongly biased towards heavy online usage (Lukawetz 2002). The author argues that greater time online improves the statistical probability of exposure to the survey, allows more time to complete it and may lead to stronger identification as an Internet user, engendering motivation to complete an online questionnaire. In this way, the principle of social validation may also be invoked when heavy Internet users respond to a web survey.

4.1.2.6 Scarcity

The use of web surveys is still in its infancy, however, and the novelty effect may be an important factor for many (Vehovar *et al.* 2002). Such novelty may invoke the principle of scarcity, where participation is understood as an unusual opportunity.

Although Sheehan's (2001) review of 31 studies using email surveys for academic purposes indicated that response rates had fallen significantly with the passing of time, the novelty value may remain for Internet newcomers. With respect to this, the figures show that British people are continuing to come online. Between 2000 and 2004, the proportion of adults reporting use of the Internet in the previous three months rose from 40% to 61% (Office for National Statistics Omnibus Survey, http://www.statistics.gov.uk).

The above has illustrated how an individual's use of and familiarity with the Internet may determine how the principles of scarcity or consistency are invoked. The novelty effect may be short-lived, however, with implications for the design of future web surveys. A survey of Internet users found that the people were mainly motivated to use the Internet because of its perceived usefulness for their job and secondarily because it was enjoyable and easy to use (Teo *et al.* 1999).

Other appeals to the principle of scarcity are illustrated in a study where telling people that had been specifically chosen to participate and giving them a deadline, increased the response rate by 8% (Porter & Witcomb 2003). Of more relevance to the present study is the appeal of the rare opportunity to answer a comprehensive set of questions about one's sexual activities. Sex is not a topic that is generally discussed in great detail and the Internet and HIV survey offers respondents an unusual chance to delve into their own sexual histories. Despite the increased amount of sex research since the AIDS epidemic highlighted the urgent need to update an understanding of sexual behaviour based on Kinsey's work of the 1940s and 50s, the number surveys of sexual behaviour undertaken in the UK remains limited.

4.1.2.7 Authority

The final principle which may influence response is authority. In this way,

> *"Attempts to persuade the sample members of the value of the survey are generally made, often supported by reference to a prestigious sponsor; good*

sponsorship is likely to be particularly effective with a mail survey" (Kalton 1983: 65).

Studies have consistently found that university sponsorship increases response rates over other organisations and the seniority or prominence of the sender also has a positive effect (Edwards *et al.* 2002). Goyder's (1987) work on survey response found that government sponsorship led to a marked increment in response. His survey on surveys also found that the most popular reason for participation volunteered by respondents was survey sponsorship, although when asked to rate a list of possible determinants, the topic of the survey was believed to be more important than university or government sponsorship.

Survey appeal has also been found to interact with sponsorship. Thus, an appeal to social utility was found to produce the best response rate for a university sponsored survey whereas an egotistic appeal induced a similar response rate for an equivalent commercially sponsored survey (Houston & Nevin 1977). The authors concluded that different sponsors could increase their response rates through the use of appropriate appeals. In the same way, the social utility of the Internet and HIV web survey is likely to be reinforced by its university sponsorship.

The above review has emphasised the varied and individual nature of the factors that are likely to have influenced motivation to participate in the Internet and HIV web survey. It demonstrates that there is likely to be a complex interaction between the inherent features of the survey – the sexual nature of the content and web mode – and the compliance principles described by Cialdini's model. The following provides an empirical examination of these factors.

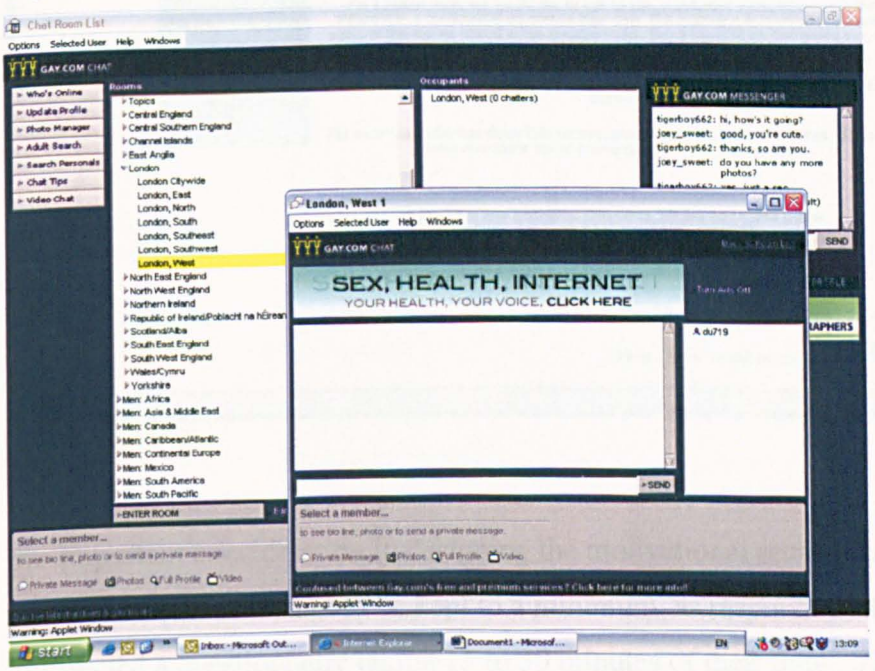
4.2 Methods

This section describes the methods used to examine motivation for participation in the Internet and HIV web survey. As described in Chapter 3, the analysis is based on primary data gathered as part of the Internet and HIV 2003 web survey and the section begins with a description of how this was undertaken. It goes on to raise the issues associated with using a survey to ask about a survey and concludes with a detailed description of the data analysis techniques used in this chapter.

4.2.1 Development of the survey instrument

In May and June 2003, respondents were recruited into the Internet and HIV web survey by clicking on pop-ups and banners which appeared in gaydar and gay.com UK chatrooms and personal profiles. The following is an example of a banner appearing at the top of the London, West chat room of gay.com:

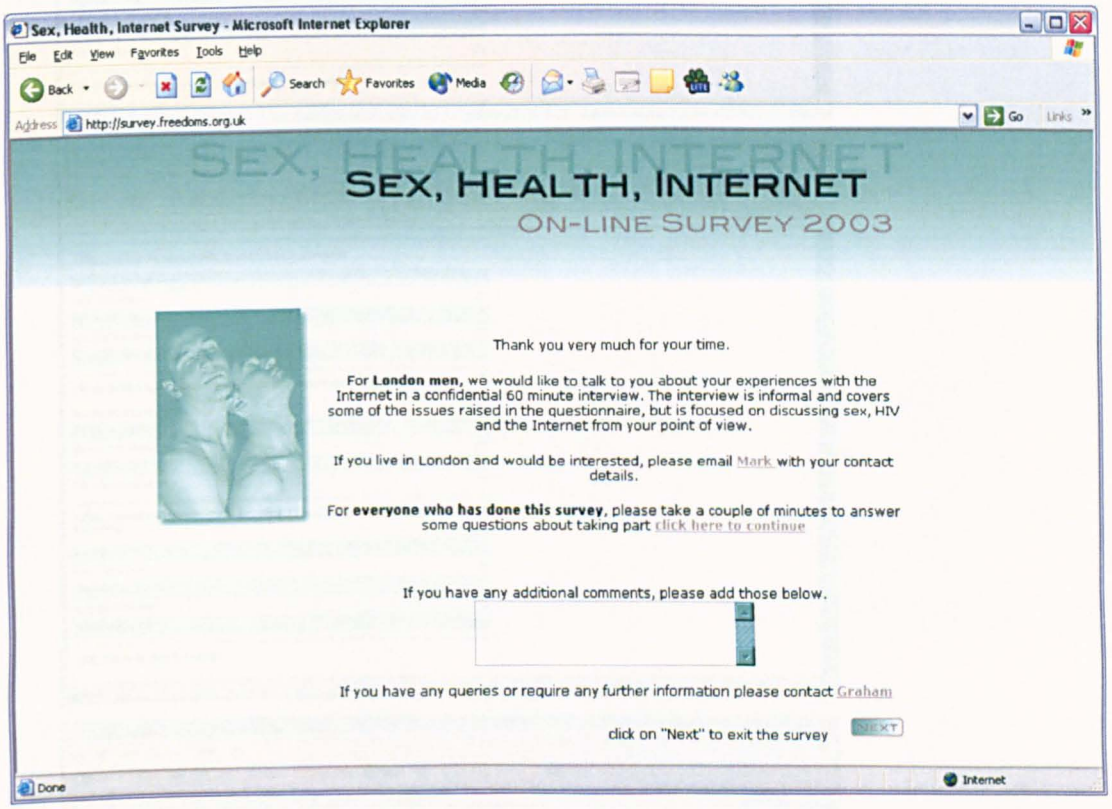
Figure 4.4: Internet and HIV banner in gay.com chatroom



When they had completed the main survey, respondents were given the opportunity to click into a further set of questions that will be referred to here as the "motivational section". Although a seamless transition to the motivational section

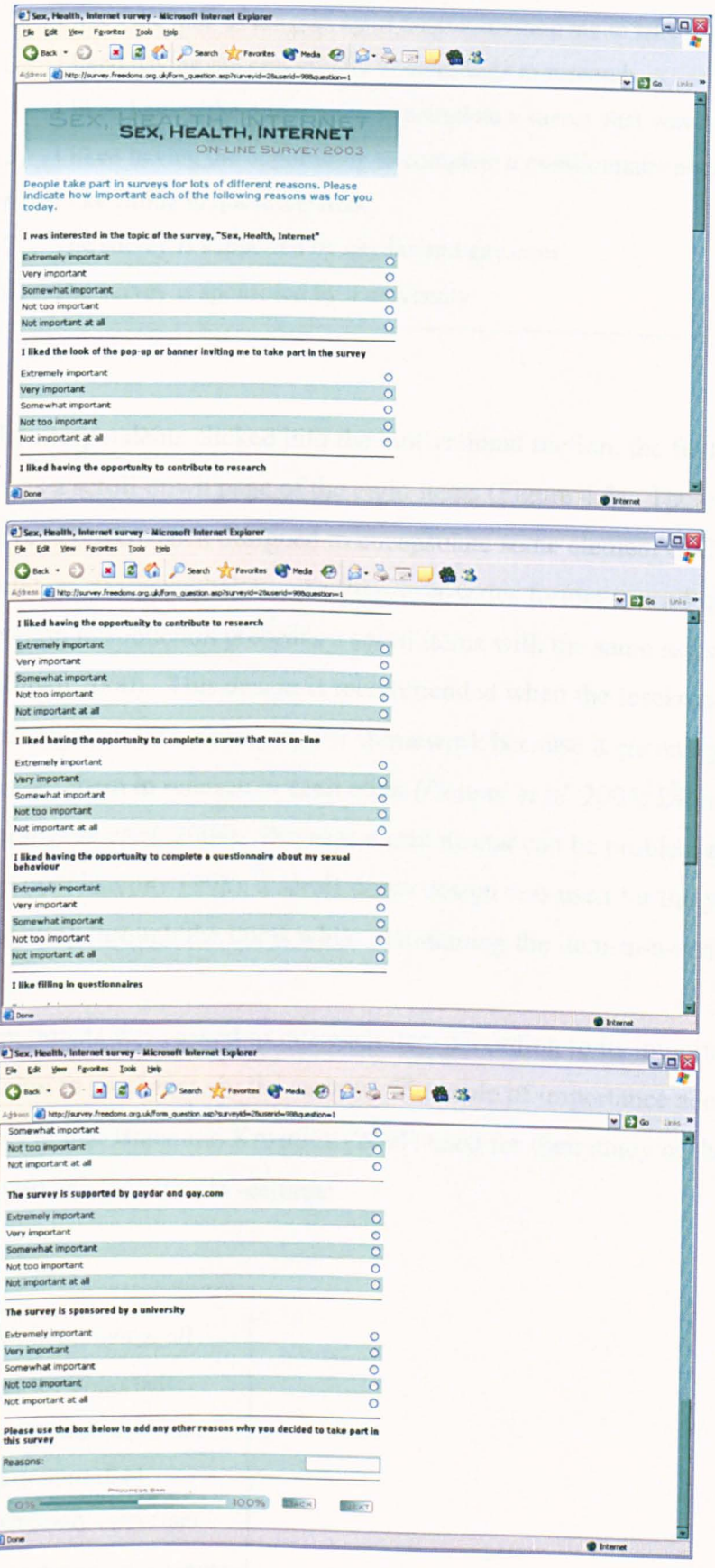
would have been desirable in order to maximise response, it was important for the invitation to participate in a one-to-one qualitative interview for the main study to appear immediately upon completion of the main survey, in order to encourage interest. On reaching the final page of the main survey (Figure 4.5), therefore, London men were asked to volunteer for a one-to-one interview and everyone who had done the survey was asked to participate in the motivational section:

Figure 4.5: Final page of Internet and HIV 2003 web survey



An important consideration in designing the motivational section was that the number of questions should be kept to a minimum, as respondents had already completed a questionnaire taking 15 to 30 minutes of their time. The number of motivational questions was therefore limited to eleven and a progress bar was included so that respondents could see their rapid progress through the questions.

Figure 4.6: Illustration of motivational items



1. I was interested in the topic of the survey, "Sex, Health, Internet"
2. I liked the look of the pop-up or banner inviting me to take part in the survey
3. I liked having the opportunity to contribute to research
4. I liked having the opportunity to complete a survey that was on-line
5. I liked having the opportunity to complete a questionnaire about my sexual behaviour
6. I like filling in questionnaires
7. The survey is supported by gaydar and gay.com
8. The survey is sponsored by a university

When respondents clicked into the motivational section, the first page that they came to was a scroll-down page of the eight items (Figure 4.6). These items are listed above and were each designed to encapsulate some element of Cialdini's compliance model, as described below. An item-in-a-series format was adopted whereby a common introduction precedes a set of items with the same response format (Dillman 2000). This design is recommended when the researcher wants to draw the items together into a comparative framework because it encourages respondents to consider them in relation to each other (Couper *et al.* 2001; Dillman 2000; Tourangeau *et al.* 2004). Because a grid format can be problematic and may lead to drop-out (Jeavons 1998), a scroll-down design was used for the purpose of navigating through the items while maintaining the item-in-a-series format.

Respondents were asked to rate each item according to its importance in their decision to participate in the survey. The scale of importance adopted was based on the scale that Bizer and Krosnick (2001) used for their study of the structure of strength-related attitude features:

Not important at all
Not too important
Somewhat important
Very important
Extremely important

In line with Krosnick's earlier work on satisficing (Krosnick 1991; Krosnick & Fabrigar 1997), whereby respondents may not expend the necessary effort to generate an optimal answer, the scale did not include a "not sure" option. Close attention was paid to item design, however, to ensure that respondents would be able to select an appropriate answer from the scale of importance. A radio button response was chosen, whereby respondents were only able to click one response option for each of the items presented.

At the bottom of this page, respondents were given the opportunity to include any other motivations that had been important for them.

Following these eight items, respondents were asked three closed questions about survey behaviour which were also related to their motivation for participation and to which they could respond "yes", "no" or "not sure":

9. In the last 12 months have you taken part in any other on-line surveys?
10. Do you know any other men who have taken part in this on-line survey?
11. Apart from this time, have you seen the pop-up or banner inviting you to take part in this survey before?

The "not sure" option was included in this case in order to provide an exhaustive range of response options, with the aim of avoiding a biased response where respondents felt unable to answer "yes" or "no" (de Vaus 1996). These questions were presented on separate pages, so that respondents could be routed to further questions according to their answers.

Two of these survey behaviour questions drew on Cialdini's model whereas the third (question 11 about having previously seen the pop-up or banner) did not, as described below.

The motivational section questions were developed over a series of stages. In the initial stage, a preliminary set of questions was developed and discussed with survey experts. These were then pre-tested with a view to improving content validity and reliability. In the first stage of piloting, copies of a questionnaire mock-up created as Word document were distributed among colleagues, friends and students. Although they were asked to imagine that they had completed the Internet and HIV web survey, pre-testing the logic of the questions about motivation for participation in the main survey was necessarily limited as none of the individuals piloted had completed the main survey and could not draw on this experience in answering the motivational items. With this limitation in mind, however, feedback was sought on the length of time for completion, the clarity of instructions, respondents' interpretation of the questions and their overall impression of the content and layout, as recommended by Fowler (1993).

In the second stage of piloting, a web version of the main survey including the motivational section was tested. Male students from the LGB Society at University College London were paid a small sum to participate in a group piloting session. Again they had not completed the survey of their own volition, so could only surmise what their answers to the motivational section might be. Such piloting is particularly important for web surveys, however, in order to minimise unanticipated behaviour of both the participants and the software when the survey goes live (Hewson *et al.* 2003).

Refinement during the piloting process resulted in the ten items listed above that were designed to tap into Cialdini's model plus an additional question (number 11) on survey response behaviour. The following describes how the items were derived. It illustrates how elements of the compliance model may be underpinned by a number of the survey items and how individual items may not map exclusively onto a single element of the model.

The intention of adopting Cialdini's framework was not to test its validity. Although such an empirical validation of the model does not appear to have been made, testing

the validity of the complex concepts underlying each of the principles would entail development of a multi-item scale with psychometric properties. The aim here was to adopt an established framework in order to derive a set of motivational items which might cover a broad range of possible motivational factors and which were operationalised with reference to the literature described earlier. These items do not therefore provide a unique or exhaustive interpretation of Cialdini's model of compliance as applied to survey participation.

Four items were designed to measure the ***inherent features*** of the activity of participating in the Internet and HIV web survey. These features relate to the act of filling in a questionnaire which was on the web and which contained questions about sex, sexual health and the Internet. The essential feature of participating in the Internet and HIV web survey was filling in the questionnaire. Although this activity is unlikely to evoke strong feelings, with over 60% of respondents reporting indifference when asked about the pleasantness of their last research experience (Nederhof 1987), the following item was designed to measure whether a general enjoyment of the activity may contribute to the proposal that survey participation is concentrated among a minority of the adult population (Bickart & Schmittlein 1999):

- I like filling in questionnaires

A second fundamental feature of filling in the questionnaire was that it was accomplished online which may also be a factor behind the initial decision to participate:

- I liked having the opportunity to complete a survey that was on-line

A third fundamental feature of the survey was its topic which, as previously described, has an important motivating force (Goyder 1987; Groves *et al.* 1992; Vehovar *et al.* 2002). When the respondent clicked into the Internet and HIV survey, its content was described as "Sex, Health, Internet" and the following item was designed to measure the impact of this description of the topic:

- I was interested in the topic of the survey, "Sex, Health, Internet"

Because the sexual nature of the content was an important feature of the topic that may have encouraged participation, its impact was isolated in the following item:

- I liked having the opportunity to complete a questionnaire about my sexual behaviour

In accordance with the ***principle of liking***, Dillman (2000) has argued that people may be more likely to complete questionnaires that they find attractive. The extent to which the initial appearance of the questionnaire encouraged respondents was therefore measured by the following:

- I liked the look of the pop-up or banner inviting me to take part in the survey

The ***principle of reciprocation*** states that people feel obliged to reciprocate when they receive something positive. Respondents may be drawn into completing the Internet and HIV web survey through the promise of the personal satisfaction associated with contributing to research and benefiting society. Contributing to research into HIV/AIDS may have an additional reciprocal benefit for men who identify themselves with a community that has been afflicted by HIV/AIDS. The importance of reciprocation was measured by the following item:

- I liked having the opportunity to contribute to research

Two items were designed to tap into the ***principle of scarcity***. The first sought to measure whether respondents who have not come across many web surveys were encouraged to participate because of a novelty effect (Vehovar *et al.* 2002):

- I liked having the opportunity to complete a survey that was on-line

in the context of

- In the last 12 months have you taken part in any other on-line surveys?

Whether the following item measures the principle of scarcity rests on the assumption that detailed discussion of sexual behaviour is not commonplace and people may therefore welcome the opportunity to answer a comprehensive set of questions about their sexual activities. The motivating force of participating in a survey specifically relating to sexual behaviour was measured by the following item:

- I liked having the opportunity to complete a questionnaire about my sexual behaviour

Previous participation in Internet surveys has been found to be an important predictor of subsequent participation (Batagelj & Vehovar 1998). It is therefore important to understand whether such surveys reach a group of respondents who are consistently willing to participate while ruling out another group who consistently refuse. The following operationalisation of the ***principle of consistency*** was used to explore whether people have a consistent propensity to participate in web surveys:

- I liked having the opportunity to complete a survey that was on-line
  in the context of
- In the last 12 months have you taken part in any other on-line surveys?

Although it proved too difficult to operationalise a succinct measure of previous survey participation, the following item is designed to tap into an enjoyment of completing questionnaires which has been associated with a general willingness to participate in research (Rogelberg *et al.* 2001):

- I like filling in questionnaires

In accordance with the ***principle of authority***, university sponsorship is consistently found to increase response rates (Edwards *et al.* 2002). This is attributed to respondents' faith in the legitimacy of surveys fielded by a trusted authority which the following item aimed to measure:

- The survey is sponsored by a university

Although gaydar and gay.com are not a traditional source of trusted authority, respondents may have a certain faith in the professionalism and discretion of their services which they may extend to a survey that is endorsed by them. The following item aimed to measure this aspect of the principle of authority:

- The survey is supported by gaydar and gay.com

At the same time, gaydar and gay.com are designed specifically for men meeting men and an association with these sites may invoke the ***principle of social validation***, to the extent that it encourages men who feel part of a gay online community to participate in a survey which is aimed at their peers:

- The survey is supported by gaydar and gay.com

At the individual level, men may also be encouraged to participate if their friends have also done so:

- Do you know any other men who have taken part in this on-line survey?

In order to examine the initial impact of the pop-up or banner on response and the issue of multiple submissions, one further item on ***survey behaviour*** was included. It asked the following question about respondents had seen the pop-up or banner before and, if so, they were directed to further questions about how they had reacted:

- Apart from this time, have you seen the pop-up or banner inviting you to take part in this survey before?

4.2.2 "A survey on surveys"[1]

Asking respondents to tell us why they participated in a survey of sexual behaviour is likely to be prone to measurement error. Groves and Couper (1998) have argued that decisions about survey participation are made at a subconscious level which might

[1] With reference to Goyder's (1986) paper "Surveys on Surveys: Limitations and Potentialities", *Public Opinion Quarterly*, vol. 50, pp. 27-41.

make it difficult for respondents to account for what prompted them to start the survey in the first place. Their answers may furthermore be subject to social desirability bias to the extent that they want to portray this decision in a positive light. As Marsh (1982) has argued, however, we should not dismiss an individual's ability to explain his own behaviour but must interpret it in the context of the following limitations:

> *"actors as such do have privileged access to their own experience but not to all the determinants of their own behaviour. However 'plausible' or 'adequate at the level of meaning' their responses are, they still may not be causally adequate. Their reasons are often much better informed and are interesting in their own right, but have not logical status over any other explanation"* (Marsh 1982: 107).

The "survey on surveys" method adopted here does not therefore provide proof as to what caused these men to start the survey but it adds to the body of knowledge surrounding such behaviour. The use of the method is also empirically supported by Goyder's (1987) finding that the results from direct questioning about survey response behaviour did not differ greatly from the analysis of behavioural data. The main cost of collecting the motivational data was furthermore subsumed under the cost of fielding the 2003 Internet and HIV web survey which therefore adds value to the Internet and HIV study while incurring little financial cost to the thesis.

4.2.3 Data analysis

A combination of SPSS 11.0 and MicroSoft Excel was used for the data analysis in this chapter, in common with the quantitative analysis for the rest of the thesis. An overall strategy for dealing with item nonresponse was adopted throughout the thesis and is detailed in Appendix 2.

The analysis in this chapter includes both bivariate and multivariate techniques. The bivariate analysis was undertaken to examine the response of subgroups of respondents to variables from the motivational section. A full list the independent variables describing the subgroups can be found in Appendix 3. They can be divided into four main types: (1) sociodemographic profile, (2) sexual orientation and

behaviour, (3) health, social and emotional, and (4) survey behaviour and Internet use. All sociodemographic items from the main survey were adopted. This includes ethnicity although it should be noted that 96.7% of respondents were white. Items on sexual orientation and behaviour, and on health, social and emotional issues were included because of their potential to influence respondents' engagement with a survey of sexual health and behaviour. Items on survey behaviour included questions from the main survey (about whether the respondent was at home during survey completion and whether he took part in the 2002 Internet and HIV web survey) and paradata that were gathered at the time of completion. Whether the respondent came from a pop-up or banner in gaydar or gay.com, the time he took to complete the survey and when he completed the survey were automatically logged with his answers to the survey questions. Items about Internet use were included to contextualise the respondents' participation in a web survey.

For the purposes of explaining the bivariate analysis, the statistical tests can be divided into those involving ordinal dependent variables and those involving dichotomous dependent variables. Different statistical tests were used for various combinations of variables in order to take account of the measurement properties of each variable.

Figure 4.7: Ordinal dependent variables

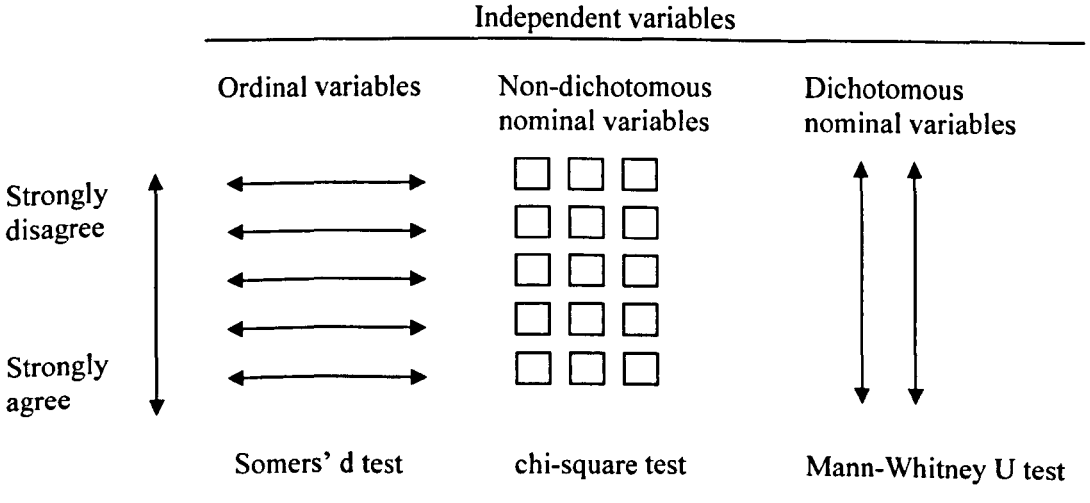
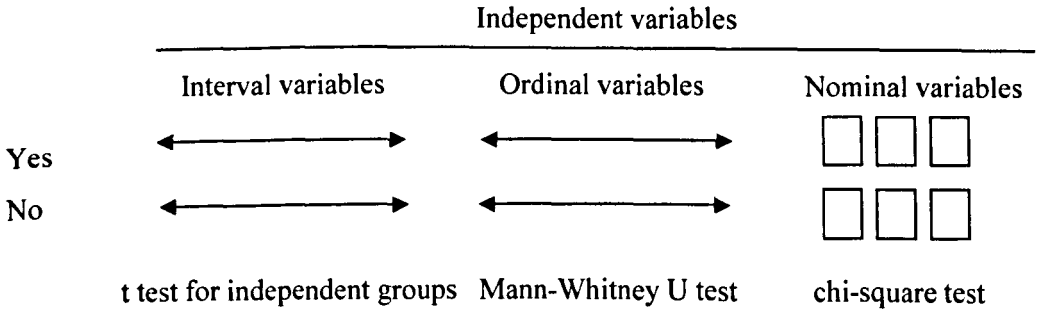


Figure 4.7 illustrates the statistical tests used with ordinal dependent variables and describes the subgroup analyses involving the eight motivational items rated on a 5 point scale of importance. Somers' d test was used in bivariate analysis involving ordinal independent variables; the chi-square test was used where independent variables were non-dichotomous nominal variables; and the Mann-Whitney U test was used where they were dichotomous nominal variables.

Figure 4.8 illustrates the statistical tests used with dichotomous dependent variables which described the yes/no response to whether respondents took part in the motivational section or agreed with the three survey behaviour items. The t test for independent groups was used with independent variables at the interval level; the Mann-Whitney U test was used where independent variables were at the ordinal level; and the chi-square test was used where they were at the nominal level.

Figure 4.8: Dichotomous dependent variables



The ***Somers' d test*** (Somers 1962 cited in Everitt 1992) is a nonparametric test that is used to examine the association between two sets of ordinal variables. Although the chi-square test may be applied in these circumstances, Somers' d takes the ordinal nature of both variables into account which provides a more fitting assessment of the data.

The ***chi-square test*** (Pearson 1904 cited in Everitt 1992) is a nonparametric test that analyses the difference between observed and expected frequencies. There is some

debate over the minimum number of expected frequencies necessary for the test to meet statistical assumptions. Although a minimum expected frequency of 5 is usually recommended (Norusis 1993), the samples used in the analysis were so large that this was rarely an issue. Everitt (1992) furthermore believes the rule to be arbitrary and finds little mathematical or empirical support for it.

The ***Mann-Whitney U test*** is a powerful test that analyses the degree of separation between samples (Pagano 2001). It is used here to capture the ordinal nature of the data. Because it is a nonparametric test, it does not carry the mathematical assumptions of the t test. Differences between two groups of respondents are illustrated in the findings by presentation of percentages of respondents in equivalent subgroups.

The ***t test for independent groups*** (Pagano 2001) is a powerful parametric test which examines differences between the mean scores of two independent groups based on assumptions that the populations from which the samples are drawn are normally distributed and their variances are equal. The t test is robust and relatively insensitive to violations of its underlying mathematical assumptions provided that $n_1 = n_2 \geq 30$ (Pagano 2001). If a pooled-variance t test is used where variances are not equal, the p level of the statistic may be in error to the extent that the sample sizes and variances are unequal (Norusis 1993). SPSS applies Levene's test of equality of variance, to examine the null hypothesis that variances are equal and where the equality of variances assumption is not met, the separate-variance t value was selected. The t test was used here to compare the responses given by age. Although samples compared were all large, they were not always of similar size and there was evidence that age was positively skewed in some samples. When the analysis was repeated with the Mann-Whitney U test it produced the same results.

These bivariate analyses were repeated for a large number of subgroup comparisons and it should be noted that when a sufficiently large number relationships are explored, they are likely to be significant on five percent of occasions by chance alone where the p value is set at 0.05 (Kish 1959).

A ***coding frame*** was drawn up to classify the motivations that respondents volunteered in response to an open question about their reasons for participating in the survey. It was used to identify how these motivations related to the inherent activity of participating in the survey, the compliance principles, the motivational dimensions that emerged from the multivariate analysis and the issue of having the time available. For example, the response "*for the benefit of other gay people*" was coded as relating to reciprocity, social validation and the altruistic dimension.

Following bivariate analysis, independent variables that were significantly associated ($p \leq 0.05$) with dependent variables were selected for ***regression modelling***, using stepwise logistic or multiple linear regression. The selection of independent variables on empirical grounds and the use of the stepwise procedure reflect the exploratory approach taken in this methodological study, as described in Chapter 3. Some further discussion is required, however, because it is often recommended that the selection of predictive variables is based on substantive theoretical rationale (Kinnear & Gray 2000) and stepwise regression has been criticised as an atheoretical approach whereby the search for patterns in the data is given over to the computer and a story is then constructed by way of explanation (Marsh 1982).

The exploratory nature of the analysis meant that there were no firm theoretical criteria for selecting predictive variables. The stepwise method was thus used to select the best set of predictive variables from those significantly associated with the dependent variable. The method takes interaction between variables into account by including and excluding explanatory variables from the model, according to the degree of correlation with the dependent variable. It controls for the effects of the other independent variables as they are entered and can therefore be used to establish the most parsimonious model. It is recommended in order to establish the optimal set of explanatory variables which Tabachnick and Fidell (2001) describe as the smallest, uncorrelated set that "covers the waterfront" with respect to the dependent variable. Stepwise regression thus accounts for the fact that explanatory variables are unlikely to be independent which was important in the present analysis where many of the explanatory variables were significantly associated. One of its

advantages therefore is that it results in the identification of just one subset of predictive variables rather than a number of equally good ones but, because of this, should not be used uncritically (Collett 2003).

Multiple linear regression examines the relationship between a set of independent variables and a dependent variable. Its primary goal is to determine the strength of the relationship and to assess the importance of each of the independent variables to that relationship (Tabachnick & Fidell 2001). It was used here to examine the relationship between the independent variables and respondent scores on the motivational dimensions revealed by the principal components analysis described below. It may also be used to compare the predictive power of a number of sets of independent variables (Hair *et al.* 1998). It uses the following equation:

$$y_i = a + b_1x_{1i} + b_2x_{2i} + \ldots b_{ki}x_{ki} + e_i$$

where i is the individual, x_1 to x_k are the independent variables, a is the intercept, b_1 to b_k are the regression coefficients and e is the residual or error term, denoting unexplained variance. The statistical significance of the coefficients is assessed using the t test.

Use of linear regression assumes that variables included in the analysis are normally distributed. The relationship between dependent and independent variables should be linear and homoscedastic (y values are normally distributed with constant variance for each x value). The residual plots are examined for evidence of departure from these assumptions and, in the present analysis, the data were considered suitable for analysis. Another assumption is that explanatory variables are independent of each other, exhibiting a low degree of multicollinearity but, as described earlier, the stepwise procedure may be used to select the best set of predictive variables from a set of correlated independent variables.

The strength of a regression model is assessed by examination of the value of the adjusted R^2, which gives the percentage of the dependent variable that is explained by the independent variables. The value of R^2 is a particularly important measure in

all regression analysis undertaken here, given that explanatory variables were not chosen on a theoretical basis. The F ratio provides a statistical test of the value of adding variables into the regression equation. The relative contribution of the explanatory variables in each model can be assessed by comparison of the standarised beta weights.

Logistic regression is used to estimate the probability of an outcome occurring or not (Norusis 1994a). Whereas linear regression models the data on the basis of the linear equation described earlier, logistic regression aims to fit data to an S shaped logistic curve (Hair *et al.* 1998). Linear regression uses the least-squares method of estimation whereby the coefficients represent the least sum of the squared distances between observed and predicted values of the dependent variable, whereas logistic regression uses the maximum-likelihood method whereby the coefficients are those that maximise the likelihood of the outcome under examination (Tabachnick & Fidell 2001). Logistic regression was used here to find the best set of predictive variables for the likelihood of respondents completing the motivational section, finding particular motivational items important, seeing the pop-up or banner before, knowing other men who had taken part and taking part in other online surveys.

Logistic regression calculates the odds ratio for each of the predictive variables which compares the probability of the outcome occurring or not occurring (Hair *et al.* 1998). The statistical significance of these coefficients is assessed using the Wald test. A good model has a high likelihood of fitting the observed results (Norusis 1994a). This is assessed using minus two times the log likelihood (-2LL) with smaller values indicating better fit. The chi-square test provides a statistical assessment of goodness-of-fit by examining the difference between -2LL for the null model where no variables are included and the final model. The classification table gives a further indication of the fit of the model by indicating the percentage of respondents that were correctly classified (Norusis 1994a).

The Cox and Snell R^2 describes the proportion of variance that is accounted for by the predictive variables in the model. As described earlier, R^2 is an important indication of whether variables used in the analysis were of any explanatory value.

One of the advantages of using logistic regression is that it is more flexible than other techniques and carries no assumptions about the distribution of the independent variables (Tabachnick & Fidell 2001). The samples used here were large enough to provide enough cases relative to predictive variables and in the unlikely event that the expected frequencies were too small for the goodness-of-fit test, it is reasonable to accept the reduced power of the analysis that may result (Tabachnick & Fidell 2001).

A further assumption is of a linear relationship between continuous predictors and the logit transformation of the outcome variable. The Box–Tidwell approach is recommended for examination of this assumption (Tabachnick & Fidell 2001). The assumption is violated when interaction terms between continuous predictors and their natural logs are included in a model and found to be statistically reliable. In the event that continuous variables in the models examined here appeared to violate the assumption, they were recoded into dichotomous variables.

In keeping with the investigative approach of the study, ***exploratory factor analysis*** was used to examine whether the motivational items could be summarized by any underlying dimensions. This is a statistical technique for identifying a small number of factors to represent the relationship between a set of many interrelated variables (Norusis 1994b). Since the main objective was not to test a theoretical model but to find a parsimonious summary of the dataset, principal components was used as the method of extraction.

Departures from assumptions of normality, homoscedasticity and linearity may be tolerated, as their effect is to diminish observed correlations upon which factor analysis is based (Hair *et al.* 1998). The size of the sample used here was more than

adequate, given that a sample of 100 is sufficient for a reliable factor analysis and a 2:1 ratio of subjects to variables is satisfactory (Kline 1994).

The first step of the analysis is to inspect the inter-item correlation matrix for the existence of relationships between variables. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy provides further evidence of the degree of intercorrelation. It enables rejection of individual items that fall below 0.5, which is "unacceptable", and assigns a value to the overall level of intercorrelation, allowing judgement of whether to proceed with the analysis. The Bartlett test of sphericity tests a null hypothesis of no correlation between variables.

Where there are no theoretical grounds in determining the number of factors to extract, the following statistical and pragmatic considerations may be taken into account. The first criterion is the eigenvalue of the factor. Since each variable has a variance of one, factors with an eigenvalue of less than one account for no more of the variance than a single variable and are excluded from the solution. Use of this criterion alone is likely to overestimate the number of factors, however (Kline 1994; Norusis 1994b). The second criterion is the scree test. This is a plot of the eigenvalues against the component number and indicates a suitable number of factors at the elbow of the curve. Although it is a subjective test, it is considered accurate to within one or two factors (Tabachnick & Fidell 2001). The third criterion is an examination of the percentage of variance accounted for by the solution. A solution that accounts for 60 percent may be considered satisfactory (Hair *et al.* 1998). The communalities after extraction describe the proportion of variance in each item that is accounted for by the solution, indicating its effectiveness for individual items.

Before a solution is accepted on the above grounds inspection is made of the component matrix. As an aid to interpretation, the axes of the original matrix are rotated. Orthogonal rotation assumes that the factors are independent and is the preferred rotational method when the goal is to reduce the data to a smaller set of uncorrelated factors for use in further analysis (Hair *et al.* 1998). The rotated matrix is then inspected and, with a sample of at least 100 respondents, factor loadings of

0.3 are considered significant (Kline 1994). Caution is advised where factors are defined by only one or two variables (Tabachnick & Fidell 2001).

Finally, a comparison of several trial solutions is recommended in order to arrive at a solution that has conceptual meaning (Hair *et al.* 1998; Rust & Golombok 1989). SPSS calculates factor scales for the final solution which are adequate for the purposes of this exploration of the data, although use of these scales may reduce generalizability of the results (Hair *et al.* 1998).

A further multivariate technique was used to explore whether respondents could be meaningfully grouped according to how they rated the motivational items. The aim of ***cluster analysis*** is to find the most appropriate way of grouping respondents on the basis of the variables included in the analysis by maximising the homogeneity of individuals within clusters and the heterogeneity between clusters (Hair *et al.* 1998; Norusis 1994b). It does this by calculating the distance between individuals and assigning them to groups on the basis of certain sorting criteria. Squared Euclidean distance is the most commonly used distance measure (Hair *et al.* 1998; Norusis 1994b) and was used in the analysis undertaken here. Although cluster analysis does not have the statistical foundations of other multivariate techniques, variables included in the analysis should not exhibit a substantial amount of multicollinearity and the inter-item correlation should be examined in this respect (Hair *et al.* 1998)

Where datasets are large, the calculations involved in cluster analysis require substantial computer resources and because the sample here contained well over 200 cases, K-means cluster analysis is recommended in order to find a solution efficiently (Norusis 1994b). The K-means cluster analysis uses a nearest centroid sorting criteria which assigns cases to the clusters where the distance between the case and the centre of the cluster is the smallest (Norusis 1994b). It requires that the number of clusters is specified and it is therefore advisable to undertake a hierarchical cluster analysis on a sub-sample of cases in order to determine the optimal number of clusters. This analysis therefore adopted the following procedure. A hierarchical cluster analysis of a randomly selected group of 10% of the whole was undertaken,

using the furthest neighbour criteria where the distance between cases and clusters is calculated on the basis of their two furthest points (Hair *et al.* 1998; Norusis 1994b). The formation of clusters is graphically illustrated by the dendrogram and this was used to determine the appropriate number of clusters by identifying a point at which the distances between clusters that are being combined is not too large. The number of clusters suggested was then specified in the K-means cluster analysis in order to examine the interpretative utility of each.

4.3 Findings

The following begins with a comparison of the men who finished the main survey and went on to complete the motivational section and those who did not. It continues with a general description of the findings from the motivational section before providing a more detailed analysis of the motivational items, the motivations that respondents volunteered and the survey behaviour items.

4.3.1 Men who completed the motivational section

Of the 4,271 men who started the 2003 Internet and HIV web survey, 2,505 men from the UK completed it and 1,366 UK men went on to complete the motivational section (54.5% of UK completers).

In order to examine the extent to which the men who completed the motivational section (n = 1,366) differed from those who did not (n = 1,139), comparisons were made between these two groups using the independent variables described earlier and listed in Appendix 3.

The men who completed the motivational items were older (mean age 34.7 vs 31.0, $p \leq 0.01$), less likely to be from an ethnic minority (3.3% vs 5.0%, $p \leq 0.05$) or to be students (11.6% vs 17.9%, $p \leq 0.01$) and more likely to have a degree or professional qualification (46.8% vs 41.5%, $p \leq 0.01$). They were more likely to live outside London (79.5% vs 75.3%, $p \leq 0.05$) and in a rural area (10.3% vs 7.8%, $p \leq 0.05$). They were more likely to identify as gay rather than bisexual (88.2% vs 80.7%, $p \leq 0.01$), were more open about their sexuality (71.2% vs 63.5%, $p \leq 0.01$) and more likely to have had sex with only men in the last year (90.1% vs 81.8%, $p \leq 0.01$). They were less likely to have a current female partner (5.0% vs 9.2%, $p \leq 0.01$) and they were more likely to have a current male partner (42.1% vs 37.9%, $p \leq 0.05$), to know gay men with HIV (44.0% vs 38.1%, $p \leq 0.01$), to be a member of a gay group (10.7% vs 7.5%, $p \leq 0.01$), to read the gay press more often (30.1% vs 25.8%, $p \leq 0.01$), visit gay venues more often (66.6% vs 61.4%, $p \leq 0.01$) and socialise more with gay men (58.4% vs 52.2%, $p \leq 0.01$).

They were more likely to have tested for HIV (56.0% vs 49.3%, $p \leq 0.01$), to be in poorer health (20.9% vs 18.7%, $p \leq 0.05$) and to have been depressed (53.3% vs 48.7%, $p \leq 0.05$).

They were more likely to have taken longer over completing the present survey (20.2% vs 13.7%, $p \leq 0.01$) and less likely to have taken part in the 2002 Internet and HIV web survey (52.0% vs 56.2%, $p \leq 0.05$).

In order to examine which of the many variables significantly associated with completing the motivational section were the most powerful predictors and their relative impact, these variables were included in a stepwise logistic regression. Table 4.1 shows the eight variables that were retained in the final model and the odds ratio or likelihood that respondents in the listed category completed the motivational section. Openness about sexual orientation and being depressed were re-coded into dichotomous variables to simplify interpretation and comparison of their odds ratio values.

Table 4.1: Variables associated with completing the motivational section

	Odds ratio
Age (continuous)	1.03 **
Survey completion time (Fast is reference)	
Medium	1.90 **
Slow	2.34 **
Sex with men only (vs men and women)	1.81 **
Lives outside London (vs in London)	1.48 **
Holds a degree or equivalent(vs not)	1.36 **
Open about sexual orientation (vs not open)	1.32 **
Depressed (vs not)	1.26 *

Note: *$p \leq 0.05$, **$p \leq 0.01$

The model accounts for 8% of the variation in whether or not the respondent completed the motivational section ($R^2 = 0.08$). It provides significant improvement over a null model where no variables are used for prediction (-2LL = 2,761.1, χ^2 = 174.1, 8df, $p \leq 0.01$), increasing the percentage of correctly classified respondents from 54.9% to 62.4%. It shows that taking longer to complete the main survey was a particularly powerful predictor of a greater likelihood of completing the motivational section. Those who completed the main survey in the slowest times were more than twice as likely to complete the motivational section as those who completed the main survey in the fastest times (OR = 2.34). As respondents got older, there was an incremental increase in the likelihood of completing the motivational section (OR = 1.03) and increased likelihood of so doing was also associated with having sex with men only (OR = 1.81), living outside London (OR = 1.48), holding a degree (OR = 1.36), being open about sexual orientation (OR = 1.32) and having been recently depressed (OR = 1.26).

4.3.2 Responses to the motivational section

Table 4.2 lists the motivational items ranked by importance attributed to them by the respondents.

Table 4.2: Level of importance attributed to motivational factors

	High (%)	Medium (%)	Low (%)	*N*
Opportunity to contribute to research	74.9	21.6	3.5	1,345
Interest in survey topic	60.4	32.8	6.9	1,327
Opportunity to complete online survey	53.0	28.9	18.1	1,351
Opportunity to complete sex survey	46.8	33.2	20.1	1,354
Supported by gaydar and gay.com	43.1	25.5	31.4	1,352
Sponsored by university	42.8	28.1	29.2	1,354
Look of pop-up or banner	24.5	31.6	43.9	1,347
Likes filling in questionnaires	16.6	27.5	55.8	1,347

Note: high importance = "extremely" or "very important"; medium importance = "somewhat important"; low importance = "not too" or "not at all important"

The opportunity to contribute to research was the most important factor in the decision to participate which three-quarters of men cited as very or extremely important. A tiny minority of less than one percent considered it to be of no importance at all. The topic of the survey, described as "Sex, Health, Internet", was the second most important factor (60.4%), followed by the opportunity to complete to an online survey (53.0%) and a survey of sexual behaviour (46.8%). Similar levels of importance were attributed to the support of gaydar and gay.com (43.1%) and a university (42.8%). Men did not consider that liking the look of the pop-up or banner was important in their decision to participate, with just one quarter citing this as important (24.5%). Liking filling in questionnaires was the least important of all the motivational factors, which 55.8% considered to be of little or no importance.

Table 4.3 lists the responses to the survey behaviour items. Nearly two thirds of the men (62.2%) said that they had responded to the pop-up or banner on the first occasion that they saw it. About half of the men had completed other online surveys and the vast majority (85.5%) did not know any other men who had taken part in this survey.

Table 4.3: Response to survey behaviour items

	Yes (%)	No (%)	Not sure (%)	*N*
Has seen pop-up or banner before	26.4	62.2	11.4	1,350
Taken part in other online surveys	48.7	40.9	10.4	1,331
Knows other men who have taken part	4.5	85.5	10.0	1,321

Table 4.4: Association between independent variables and motivational items

	Opportunity to contribute to research	Interest in survey topic	Opportunity to complete online survey	Opportunity to complete sex survey	Supported by gaydar and gay.com	Sponsored by university	Look of pop-up or banner	Likes filling in questionnaires
Sociodemographic profile								
Getting older [1]	pos							neg
White (vs minority ethnic) [2]								
Being a student [2]	neg				neg			
Being in work [2]								
Not working [2]								
Higher social class [1]					neg	pos		
Holding a degree or equivalent [2]				neg	neg	pos		neg
Living in London (vs elsewhere in the UK) [2]					neg			
Living in an urban area [2]						pos	pos	
Born in the UK [2]					pos			
Sexual orientation and sexual behaviour								
Sexual orientation								
Identifying as gay (vs bisexual) [2]	pos	pos				pos		
Being more open about sexual orientation [1]	pos	pos					neg	
Sexual behaviour								
Sex with men only (vs men and women) [2]								
Seeking cybersex [2]			pos	pos			pos	
Looking at online porn [2]		neg						
Having high risk sex [3]								
Health, social and emotional								
Health								
Being HIV positive [2]							neg	
Being HIV negative [2]	pos	pos						
Having tested for HIV [2]	pos	pos						
Having better health [1]		pos		pos				
Social								
Having a female partner [2]								
Having a male partner [2]								
Knowing men with HIV [2]	pos				neg			neg
Being member of gay group [2]						pos		
Reading gay press [1]	pos	pos			pos	pos	pos	pos
Socialising with gay men [1]		pos						
Visiting gay venues [1]					pos			
Emotional								
Being depressed [1]						neg		
Having suicidal thoughts [1]						neg		pos
Being lonely [1]				pos			pos	pos
Survey behaviour and Internet use								
Survey behaviour								
Being at home during survey completion [2]					pos			
Starting from gaydar (vs gay.com) [2]			pos					
Having taken part in 2002 online survey [3]			neg		pos			
Longer survey completion time [1]								
Taking part in the small hours [2]								
Taking part in the morning [2]								
Taking part in the afternoon [2]								
Taking part in the evening [2]								
Later survey participation (1st to 5th week) [1]						neg		
Internet use								
Increasing hours on Internet [1]			pos			neg		pos
Increasing frequency of online cruising [1]					pos			
Recency of starting online cruising [1]								
Increasing hours online cruising [1]				pos	pos	neg		

p < or = 0.01

0.01 < p < or = 0.05

pos positive association between independent and dependent variables

neg negative association between independent and dependent variables

[1] Ordinal variable (Somers' d test)

[2] Dichotomous variable (Mann-Whitney U test)

[3] Non-dichotomous nominal variable (chi-square test)

4.3.3 Analysis of the motivational items

Table 4.4 is a summary of the association between subgroups of respondents defined by the independent variables (as described above) and the importance that they attributed to the motivational items. The left hand column lists the independent variables under the four types described earlier. Significant associations (at the 0.01 and 0.05 level) between these variables and the eight motivational items are highlighted. The table shows whether the independent variables were positively or negatively associated with finding the motivational item important.

As the practice of combining categories for analysis is not recommended (Everitt 1992), all bivariate analyses for the motivational items were carried out using the full five-point scale of importance. The percentages presented below describe the combined response for the "extremely important" and "very important" options in order to illustrate what was "important" to the respondents. Where the independent variable is at the ordinal level, the percentages reported show the importance of the motivational item to the subgroups described by either extreme of the ordinal scale, unless otherwise stated.

The following is a summary of differences between subgroups that were statistically significant at the 0.01 and 0.05 level. It should be noted that the percentage differences shown are generally not very great and that the significant relationships found between independent and dependent variables were not very strong. For example, Somers' d measures the association between ordinal variables such that values close to an absolute value of 1 indicate a strong relationship and those close to 0 indicate little or no relationship. Values for Somers' d in the following analyses ranged from 0.042 to 0.093, indicating that the relationships were all quite weak.

Opportunity to contribute to research

Gay men were more likely than bisexual men to believe that contributing to research motivated their participation (76.4% vs 64.3%, $p \leq 0.01$), as were men who were more open about their sexual orientation (77.3% vs 61.2%, $p \leq 0.01$). Men who were gay were also more open about their sexuality ($p \leq 0.01$).

Contributing to research was more important for HIV negative men (77.8%, $p \leq 0.01$) and less important for "never tested" men (71.2%, $p \leq 0.01$). It was also more important for men who knew other men with HIV compared to those who did not (76.7% vs 73.4%, $p \leq 0.01$), for those who read the gay press (79.7% vs 70.2%, $p \leq 0.01$) and who visited gay venues (76.9% vs 68.6%, $p \leq 0.05$).

Interest in survey topic

Older men were more likely to cite the importance of the topic of the survey (68.2% vs 52.9%, $p \leq 0.01$). In relation to this, students were less likely than non-students to believe it was important (48.7% vs 61.9%, $p \leq 0.05$). Gay men were more likely than bisexual men to find the survey topic important (61.3% vs 52.2%, $p \leq 0.01$) as were men who were more open about their sexuality (64.3% vs 51.1%, $p \leq 0.01$). Men who had looked at online porn were less likely to find the topic of the survey important compared to men who had not used online porn (56.9% vs 66.5%, $p \leq 0.01$).

Men who were HIV negative were more likely to be influenced by the survey topic (66.8%, $p \leq 0.01$) than men who were HIV positive (62.2%) and men who had never tested for HIV (52.9%, $p \leq 0.01$). Men who were more likely to cite this reason were also in better health (65.2% in excellent health vs 55.0% in fair health, $p \leq 0.05$), more likely to read the gay press (64.1% vs 55.4%, $p \leq 0.05$) and more socially involved in the gay community (47.5% at point 1, the bottom of the gay community scale, rising to 68.1% at point 5 and dropping back to 52.9% at point 7, the top of the scale, $p \leq 0.05$).

Opportunity to complete online survey

Men who had sought cybersex on the Internet were more likely to have appreciated the opportunity to fill in an online survey (58.4% vs 48.8%, $p \leq 0.01$).

Men who had taken part in the 2002 web survey were more likely to consider this as important than those who had not (60.1% vs 48.9%, $p \leq 0.05$), as were those who

completed the survey more speedily (62.0% vs 51.4%, $p \leq 0.01$) and those who spent longer on the Internet per week (57.0% vs 47.2%, $p \leq 0.01$).

Half of the men had taken part in other online surveys and those who had done so were more likely to believe that the opportunity to take part in an online survey was a motivating factor than men who had not (56.4% vs 48.1%, $p \leq 0.05$). Men who had taken part in other online surveys had participated in an average of 2 such surveys (median = 2; mean = 4.05, s.d. 7.9).

Opportunity to complete sex survey

Men who held a degree or professional qualification were less likely to cite this motivation than those who did not (44.6% vs 48.7%, $p \leq 0.05$). Men who had sought cybersex were more likely to consider that completing a survey of sexual behaviour was important in their decision to participate than those who had not (50.7% vs 43.7%, $p \leq 0.01$).

Men who were in better health were generally more likely to believe that filling in a survey of sexual behaviour was important (53.6% in excellent health vs 41.0% in fair health, but rising to 50.8% in poor health, $p \leq 0.05$). Men who had been lonely were more likely to believe this was a motivating force (52.0% vs 43.8%, $p \leq 0.01$).

An increasing amount of Internet cruising per week was associated with increased likelihood of citing the importance of this motivation (49.0% vs 43.9%, $p \leq 0.05$).

Supported by gaydar and gay.com

Students were less likely than non-students to consider that support by gaydar and gay.com was important (35.8% vs 44.1%, $p \leq 0.05$), as were men who had obtained a degree or professional qualification (38.3% vs 47.3%, $p \leq 0.01$) and men who lived in London (32.7% vs 45.8%, $p \leq 0.01$). Lower social class was associated with finding this motivational factor important (50.0% in Class IV vs 28.4% in Class I, $p \leq 0.01$) as was being born in the UK (44.2% vs 33.3%, $p \leq 0.01$).

Men who knew gay men with HIV were less likely to consider that support by gaydar and gay.com was important than men who did not (41.1% vs 44.4%, $p \leq 0.01$), whereas were men who read the gay press were more likely to believe that their support was important (45.5% vs 35.1%, $p \leq 0.05$), as were men who visited gay venues more often (55.6% of those who visited every day vs 33.1% of those who visited once a year, $p \leq 0.05$).

Men who entered the survey from gaydar were more likely to consider the support of gaydar and gay.com as important (45.1% vs 38.5%, $p \leq 0.05$) as were those who took longer completing the main survey (48.4% vs 40.2%, $p \leq 0.05$). Men who logged on to cruise more frequently were more likely to cite this as an important reason (46.4% vs 27.2%, $p \leq 0.01$), as were those who spent more time online cruising per week (45.7% vs 40.3%, $p \leq 0.05$).

Sponsored by university

Men who held a degree or professional qualification were more likely to cite this motivation (45.6% vs 40.3%, $p \leq 0.05$), as were men of a higher social class (48.1% vs 33.3%, $p \leq 0.05$) and those who lived in an urban area (44.1% vs 31.4%, $p \leq 0.05$). Gay men were more likely to find the support of a university important than bisexual men (44.4% vs 30.8%, $p \leq 0.01$).

Men who were members of a gay group were more likely to cite this reason (54.8% vs 41.5%, $p \leq 0.01$), as were men who read the gay press more often (45.0% vs 37.1%, $p \leq 0.05$) and men who had *not* felt depressed (44.6% vs 37.6%, $p \leq 0.05$). The negative association between having had suicidal thoughts ($p \leq 0.05$) and the importance of university support did not reveal a clearly interpretable linear relationship.

Men who participated in later weeks were less likely to consider university support as important (40.6% vs 45.6%, $p \leq 0.05$), as were those who spent more time per week on the Internet (37.7% vs 46.7%, $p \leq 0.01$) and those who spent more time online cruising (37.1% vs 44.6%, $p \leq 0.05$)

Look of pop-up or banner

Men who lived in an urban area were more likely to cite the importance of the look of the pop-up or banner in their decision to participate (25.3% vs 18.6%, $p \leq 0.05$).

Men who were more open about their sexuality were less likely to have found the look of the pop-up or banner important (22.6% vs 28.3%, $p \leq 0.05$) whereas men who had sought cybersex were more likely to find it important (26.9% vs 22.5%, $p \leq 0.01$).

HIV positive men were less likely to believe the pop-up or banner was important than HIV negative and never tested men (17.2% vs 25.2%, $p \leq 0.05$), whereas men who read the gay press were more likely to have found it important (27.9% vs 21.8%, $p \leq 0.05$), as were men who had been lonely (28.2% vs 21.8%, $p \leq 0.01$).

Likes filling in questionnaires

Younger respondents were more likely to consider that liking filling in surveys was an important factor in their decisions to participate (23.4% vs 12.8%, $p \leq 0.01$), as were men who did not hold a degree or professional qualification (20.4% vs 12.3%, $p \leq 0.01$).

Men who knew men who are HIV positive were less likely to cite this reason (14.5% vs 18.4%, $p \leq 0.01$) whereas men who read the gay press were more likely (20.5% vs 14.8%, $p \leq 0.05$), as were men who had experienced suicidal thoughts (28.0% vs 15.6%, $p \leq 0.05$) and those who had been lonely (22.8% vs 12.2%, $p \leq 0.01$) compared to those who had not.

Men who spent more time on the Internet were more likely to cite a liking for filling in questionnaires as an important motivational factor (21.0% vs 13.9%, $p \leq 0.01$).

Inter-item correlation

The inter-item correlation was examined to explore the independence of the motivational items. Table 4.5 (on the following page) shows that there is a modest

degree of inter-item correlation between the motivational items, with correlations ranging from 0.07 ("liked filling in questionnaires" and "was interested in topic of survey") to 0.53 ("liked completing a survey on sexual behaviour" and "liked completing an online survey"), with an average inter-item correlation of 0.31 and Cronbach's alpha of 0.78.

Table 4.5: Correlation between motivational variables

	Opportunity to contribute to research	Interest in survey topic	Opportunity to complete online survey	Opportunity to complete sex survey	Supported by gaydar and gay.com	Sponsored by university	Look of pop-up or banner	Likes filling in questionnaires
Opportunity to contribute to research	1.00							
Interest in survey topic	0.45	1.00						
Opportunity to complete online survey	0.45	0.36	1.00					
Opportunity to complete sex survey	0.41	0.42	0.53	1.00				
Supported by gaydar and gay.com	0.24	0.24	0.36	0.37	1.00			
Sponsored by university	0.23	0.24	0.23	0.23	0.38	1.00		
Look of pop-up or banner	0.26	0.24	0.44	0.35	0.35	0.23	1.00	
Likes filling in questionnaires	0.16	0.07	0.38	0.36	0.26	0.12	0.32	1.00

Note: average inter-item correlation = 0.31; Cronbach's alpha = 0.78

This level of correlation indicates that a multivariate analysis may successfully reveal underlying dimensions to these items. The lack of high inter-item correlation is evidence that items are not duplicating one another and should all be included in the analysis. It shows that although the operationalisation of Cialdini's compliance principles was not unique or exhaustive, it served its purpose which was to provide a set of items measuring distinct aspects of motivation for participation.

Multivariate analysis of motivational items

Logistic regression was undertaken to explore the power of the independent variables to explain whether each of the eight motivational items was "important" or "not important". This analysis revealed that the independent variables had little predictive power, with values of R^2 ranging from .00 to .03. The details of this analysis are not reported here because the values of R^2 support the earlier finding that the independent variables revealed only limited differentiation with regard to how respondents rated the importance of the motivational variables.

The inter-item correlations described above suggest that the motivational items might be effectively summarised using a multivariate approach. Exploratory factor analysis with principal components extraction was used to investigate whether there were any underlying dimensions to the eight items that respondents had rated on the scale of importance.

The generally modest degree of correlation between items was supported by a "meritorious" degree of intercorrelation (KMO = 0.828) (Kaiser 1974 cited in Norusis 1994b), which the Bartlett test (Hair *et al.* 1998) found to be significant (χ^2 = 2318.147, 28df, $p \leq 0.01$). The lowest KMO measure for individual items was 0.791, which borders on what Kaiser characterises as "meritorious". All items were therefore maintained in the analysis.

A two factor solution appeared reasonable, based on acceptance of factors with an eigenvalue of greater than one and inspection of the scree plot. This solution accounted for 53.3% of the variance and, although it explained only 31.4% of the variance in the university support item and 38.8% of the variance in the gaydar and gay.com support item, it accounted for between 47.0% and 68.6% of the variance in the remaining items.

Following orthogonal rotation, all items loaded significantly onto at least one of the dimensions which were labelled altruistic and egotistical according to their properties. Table 4.6 (on the following page) shows that the conceptual cohesion of

this solution is persuasive. An interest in the topic of the survey, the opportunity to contribute to research and university sponsorship load onto the first dimension which describes an altruistic motivation to take part in academic research. Liking filling in questionnaires, the opportunity to complete an online survey, the look of the pop-up or banner, the opportunity to complete a sex survey and the support of gaydar and gay.com loaded onto the second dimension which describes an egotistical motivational orientation, concerned with the personal attraction and benefits of participation. Three items loaded significantly onto both dimensions (opportunity to complete online survey, opportunity to complete sex survey, supported by gaydar and gay.com) but they do not load equally onto both. Whereas the loading and nature of the opportunity to complete an online survey and a sex survey can be clearly subsumed into the egotistical dimension, the loading of the support of gaydar and gay.com is less clear-cut and illustrates how this item is not particularly well described by either dimension.

Table 4.6: Loading of motivational factors onto egotistical and altruistic dimensions

	Altruistic	Egotistical
Interest in survey topic	**0.81**	
Opportunity to contribute to research	**0.72**	
Sponsored by university	**0.53**	
Likes filling in questionnaires		**0.82**
Opportunity to complete online survey	0.40	**0.67**
Look of pop-up or banner		**0.65**
Opportunity to complete sex survey	0.45	**0.61**
Supported by gaydar and gay.com	0.34	**0.52**

In order to examine the relative importance of altruistic and egotistical motivations to the individual, an altruistic and egotistical score was derived for each respondent by taking the mean of his scores for the variables loading onto either dimension. The

derived scores were rated on a scale from 1 to 5 where 1 was "not at all important" and 5 was "extremely important". The overall mean score was higher for altruism than egotistism (3.67 vs 3.03), showing that altruistic concerns were more important to the respondents on average than egotistical concerns.

A K-means cluster analysis using nearest centroid sorting criteria was undertaken to explore whether respondents could be grouped according to how they rated the motivational items. Hierarchical cluster analysis was used on a randomly selected sub-sample of respondents in order to determine an appropriate number of clusters. Inspection of the dendrogram indicated that a two or three cluster solution would be appropriate and the K-means cluster analysis was used to explore the interpretive utility of two and three cluster solutions for the whole sample.

The two-cluster solution appeared to group respondents according to their propensity to rate items as either important or unimportant. It grouped respondents into a first cluster which tended to rate all items "very important", apart from liking the look of the pop-up or banner and liking filling in questionnaires, which were rated "somewhat important"; and a second cluster which tended to rate all items as either "not too important" or "somewhat important", apart from the opportunity to contribute which was rated as "very important".

Table 4.7: Scores on the motivational items for the three-cluster solution

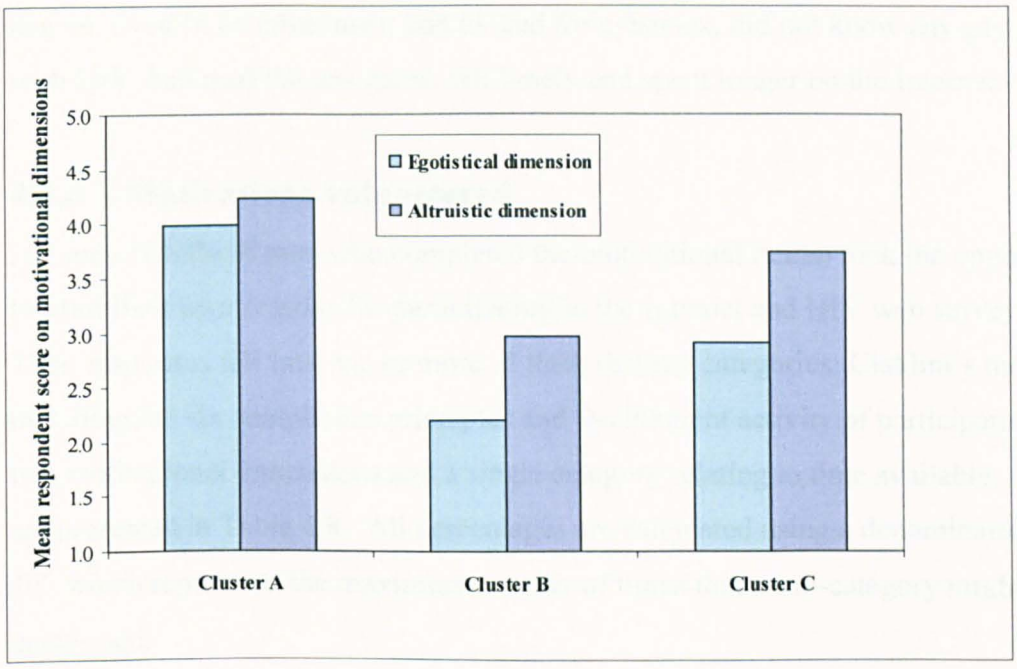
	Opportunity to contribute to research	Interest in survey topic	Opportunity to complete online survey	Opportunity to complete sex survey	Supported by gaydar and gay.com	Sponsored by university	Look of pop-up or banner	Likes filling in questionnaires
Cluster A	Extremely Important (5)	Very Important (4)	Very Important (4)	Very Important (4)	Very Important (4)	Very Important (4)	Very Important (4)	Somewhat Important (3)
Cluster B	Very Important (4)	Somewhat Important (3)	Somewhat Important (3)	Somewhat Important (3)	Not too Important (2)	Not too Important (2)	Not too Important (2)	Not too Important (2)
Cluster C	Very Important (4)	Very Important (4)	Somewhat Important (3)	Somewhat Important (3)	Somewhat Important (3)	Very Important (4)	Somewhat Important (3)	Not too Important (2)

Table 4.7 shows the three-cluster solution which revealed a more discriminating method of grouping respondents. The table shows how members of the three clusters, A, B and C, tended to rate each of the motivational items on a scale of importance from 1 to 5 where 1 was "not at all important" and 5 was "extremely important". It shows that Cluster A describes a group of respondents who tend to rate the items provided as more important in their decision to participate and Cluster B describes a group of respondents who tend to rate the items as less important. The three-cluster solution defines a third group, cluster C, which rates the survey topic, contribution to research and university support as "very important", liking filling in questionnaires as "not too important" and the other items as "somewhat important". This cluster is similar to the altruistic dimension revealed by the principal components analysis. It accounts for 40.6% of respondents, with 30.4% in cluster A and 29.0% in cluster B.

Both the two and three-cluster solutions provide further support for the pervasive importance of contributing to research as a motivational force, no matter how respondents rate other items. Bivariate analyses to examine of the association between the sample subgroups and cluster membership revealed very few significant findings.

In order to examine the relative importance of the altruistic and egotistical dimensions for respondents grouped into the three clusters, the mean of the derived scores on the altruistic and egotistical dimensions (as described above) was calculated for each of the clusters. The results are shown in Figure 4.9 (on the following page), which reinforces the idea that the cluster analysis divides the sample into groups with a propensity to rate items as important to a greater or lesser degree. It shows how respondents across all three clusters were motivated by both altruistic and egotistical considerations but that the altruistic dimension is more important for members of each of these groups. The response set adopted by Cluster A results in the least discrimination between items comprising the altruistic and egotistical dimensions whereas members of Clusters B and C show greater discrimination.

Figure 4.9: Importance of motivational dimensions to respondents in cluster groups



Stepwise multiple linear regression was used to explore the relationship between the independent variables and respondent scores on the two motivational dimensions of egotism and altruism. Variables were included that were significantly related to any of the items that composed the dimension. Whereas previous analysis used respondent scores calculated from the mean of the variables loading on to each dimension, in this case the scores were calculated using the regression-method. The mean method was used previously so that scores could be easily interpreted and graphically represented on a scale from 1 to 5. The regression method standardises scores around zero making its interpretation more complex but it is used in the regression analysis because it provides an exact measure of respondent scores derived from the principal components analysis (Norusis 1994b). The large number of variables included in this multivariate analysis resulted in a considerable amount of missing data which was dealt with through mean substitution.

Although variables included in the model had little predictive power for the altruistic dimension ($R^2 = 0.03$, $F = 13.44$, $p \leq 0.01$), the analysis indicated that altruism was likely to increase among gay, older respondents who had tested for HIV. The variables had slightly more predictive power for the egotistical dimension ($R^2 = 0.05$,

$F = 10.52$, $p \leq 0.01$) which was likely to be higher among men who did not have a degree, lived in an urban area, had looked for cybersex, did not know any gay men with HIV, had read the gay press, felt lonely and spent longer on the Internet.

4.3.4 Motivations volunteered

203 men (14.9% of men who completed the motivational items) took the opportunity to offer their own reasons for participating in the Internet and HIV web survey. Their responses fell into one or more of three distinct categories: Cialdini's model including the six compliance principles and the inherent activity of participation, the two motivational dimensions and a single category relating to time available. They are presented in Table 4.8. All percentages are calculated using a denominator of 203 which represents the maximum number of times that a sub-category might be mentioned.

Table 4.8: Motivations for participation volunteered by respondents

Coding category	Motivation volunteered	Number of responses (percentage)
Compliance model	Reciprocation	84 (41.4%)
	Inherent activity	41 (20.2%)
	Social validation	23 (11.3%)
	Consistency	6 (3.0%)
	Liking	1 (0.5%)
	Authority	0 (0.0%)
	Scarcity	0 (0.0%)
Motivational dimensions	Altruistic	89 (43.8%)
	Egotistical	67 (33.0%)
Time	Time available	26 (12.8%)

In corroboration with earlier findings, the most popular motivation volunteered was reciprocity (41.4%), which generally referred to contributing to research. The inherent activity of survey participation attracted one fifth of respondents who tended

to be bored or curious about the content of the questionnaire. Social validation was highlighted by 11.3% of men who participated, for example, "*to help other gay men indirectly*". Other elements of the compliance model received little or no mention. The motivations that were volunteered also support the earlier findings with regard to the importance of altruistic and egotistical concerns, with 43.3% coded as altruistic and 33.0% as egotistical. 12.8% of respondents said that having time on their hands was a key factor in their decision to participate.

4.3.5 Analysis of the survey behaviour items

The three survey behaviour variables were re-coded into dichotomous variables, whereby respondents who had answered "not sure" were removed from the analysis, leaving just "yes" and "no" responses. Although the "not sure" response was included to provide an exhaustive list of response options, it did not contribute to the interpretation of data and was removed on this basis. Table 4.9 (on the following page) is a summary of the association between the subgroups defined by the independent variables and responses to the survey behaviour items. Associations significant at the 0.01 and 0.05 level are highlighted and whether the independent variables were positively or negatively associated with the survey behaviour items is shown.

Has completed other online surveys

Men who held a degree or professional qualification were more likely to have completed other online surveys (57.4% vs 51.7%, $p \leq 0.05$) as were those from higher social classes, with men from class II the most likely and men from class V the least likely (55.9% vs 28.6%, $p \leq 0.05$). Men who had looked at online porn were more likely to have taken part in other online surveys than those who had not (57.6% vs 48.7%, $p \leq 0.01$). Men who took part in the 2002 web survey were more likely to have taken part in other online surveys (73.5% vs 37.8%, $p \leq 0.01$), as were men who entered the survey from gaydar (57.3% vs 47.7%, $p \leq 0.05$), those who completed the survey during the first week that it was online (61.9% vs 48.4%, $p \leq 0.01$), those who had been online cruising for longer (59.5% vs 52.7%, $p \leq 0.01$) and those who spent longer on the Internet per week (increasing from 43.3% of men who

Table 4.9: Association between independent variables and survey behaviour items

	Has completed other online surveys	Knows other men who have taken part	Has seen pop-up or banner before
Sociodemographic profile			
Getting older [1]			
White (vs minority ethnic) [2]			neg
Being a student [2]			
Being in work [2]			
Not working [2]			
Higher social class [3]	pos		neg
Holding a degree or equivalent [2]	pos		
Living in London (vs elsewhere in the UK) [2]			
Living in an urban area [2]			
Born in the UK [2]			
Sexual orientation and sexual behaviour			
Sexual orientation			
Identifying as gay (vs bisexual) [2]			
Being more open about sexual orientation [3]		pos	
Sexual behaviour			
Sex with men only (vs men and women) [2]			
Seeking cybersex [2]			
Looking at online porn [2]	pos		
Having high risk sex [2]			
Health, social and emotional			
Health			
Being HIV positive [2]			
Being HIV negative [2]			
Having tested for HIV [2]		pos	
Having better health [3]			
Social			
Having a female partner [2]			
Having a male partner [2]		pos	
Knowing men with HIV [2]		pos	
Being member of gay group [2]			
Reading gay press [3]		pos	
Socialising with gay men [3]		pos	
Visiting gay venues [3]		pos	
Emotional			
Being depressed [3]			
Having suicidal thoughts [3]			
Being lonely [3]			
Survey behaviour and Internet use			
Survey behaviour			
Being at home during survey completion [2]			
Starting from gaydar (vs gay.com) [2]	pos		
Having taken part in 2002 online survey [2]	pos	pos	pos
Longer survey completion time [3]			pos
Taking part in the small hours [2]		neg	
Taking part in the morning [2]			pos
Taking part in the afternoon [2]			
Taking part in the evening [2]			
Later survey participation (1st to 5th week) [3]	neg		pos
Internet use			
Increasing hours on Internet [3]	pos		
Increasing frequency of online cruising [3]			pos
Recency of starting online cruising [3]	neg		neg
Increasing hours online cruising [3]			pos

p < or = 0.01 | 0.01 < p < or = 0.05

pos / neg positive or negative association between independent and dependent variables

[1] Interval variable (t test)
[2] Nominal variable (chi-square test)
[3] Ordinal variable (Mann-Whitney U test)

spent up to 5 hours to a peak of 67.6% of those who spent 21-25 hours, decreasing to 55.6% of men who spent at least 31 hours, $p \leq 0.01$).

As previously stated, men who had taken part in other online surveys were more likely to believe that the opportunity to take part in an online survey was a motivating factor than men who had not and had taken part in an average of 2 such surveys.

Knows other men who have taken part

Men who were more open about their sexuality were more likely to know other men who had taken part (6.5% vs 0.0%, $p \leq 0.01$) as were men who had tested for HIV (6.4% vs 3.3%, $p \leq 0.05$), those who had a current male partner (6.7% vs 3.9%, $p \leq 0.05$) and those who knew gay men with HIV (7.0% vs 3.6%, $p \leq 0.01$). Knowing other men who had taken part was also associated with socialising more with gay men (6.3% vs 1.3%, $p \leq 0.01$), visiting gay venues more often (8.7% vs 3.1%, $p \leq 0.01$) and reading the gay press (6.4% vs 2.5%, $p \leq 0.05$).

Men who took part in the 2002 web survey were more likely to know other men who had taken part (8.3% vs 3.2%, $p \leq 0.01$), whereas men who took part in the small hours were less likely (0.0% vs 5.5%, $p \leq 0.05$).

Has seen the pop-up or banner before

There was a significant association between social class and whether respondents completed the survey the first time they saw it, illustrated by the finding that manual workers were more likely to have seen the pop-up or banner before than professionals (42.9% vs 23.1%, $p \leq 0.05$). Minority ethnic men were also more likely to have seen the pop-up or banner before (52.5% vs 29.2%, $p \leq 0.05$).

Men who had taken part in the 2002 web survey were more likely to have seen the pop-up or banner before than men who had not (41.4% vs 22.7%, $p \leq 0.01$), as were men who took part in later weeks (40.4% vs 23.8%, $p \leq 0.01$). Men who took part in the morning were more likely to have seen the pop-up or banner before (36.5% vs

27.9%, $p \leq 0.01$), as were men who completed the survey more slowly (33.9% vs 25.5%, $p \leq 0.05$).

Spending more time online cruising per week (35.2% vs 26.2%, $p \leq 0.01$), logging on to cruise more often (33.2% vs 22.6%, $p \leq 0.01$) and having used the Internet for online cruising over a longer period (34.0% vs 20.8%, $p \leq 0.01$) were all associated with a greater likelihood of having seen the pop-up or banner before.

Multivariate analysis of survey behaviour items

Logistic regression was used to explore the relationship between the independent variables and the survey behaviour dependent variables. The independent variables that were significantly associated ($p \leq 0.05$) with the survey behaviour items in the bivariate analysis were included in the models.

Table 4.10: Variables associated with taking part in other online surveys

	Odds ratio
Started from gaydar (vs gay.com)	1.34 *
Extensive use of Internet (vs limited use)	1.50 **
Completion of 2002 web survey (Yes is reference)	
No	0.22 **
Not sure	1.27
Week of survey completion	0.89 *

Note: *$p \leq 0.05$, **$p \leq 0.01$

Variables used to predict whether or not respondents had seen the pop-up or banner before accounted for a small percentage of the variation in this item ($R^2 = 0.06$) and was not substantially better than the null model, increasing the percentage of correctly classified respondents from 70.4% to 71.4% (-2LL = 1,131.8, $\chi^2 = 57.91$, 6df, $p \leq 0.01$). The model tested to predict whether respondents knew anyone else who had taken part in the survey failed to converge. However, a reasonable

proportion of the variance in whether the respondent had taken part in other online surveys ($R^2 = 0.15$) was accounted for by the model that is presented in Table 4.10. The final model increased the percentage of correctly classified respondents from 55.9% to 68.3% (-2LL = 1,253.9, $\chi^2 = 170.7$, 5df, $p \leq 0.01$).

The model shows that starting from gaydar increases the likelihood of having taken part in other online surveys (OR = 1.34) and that extensive use of the Internet in the past week (more than 10 hours) is also associated with a increased likelihood of taking part in other online surveys (OR = 1.50). Not having taken part in the 2002 Internet and HIV web survey is associated with a considerable decrease in the likelihood of taking part in other online surveys compared to those respondents who took part in 2002 (OR = 0.22), whereas being unsure about this is not a significant predictor of online survey participation. Taking part in later weeks was associated with an incremental decrease in the likelihood of taking part in other online surveys (OR = 0.89).

4.4 Discussion of findings

Adopting Cialdini's compliance principles as a framework for operationalising the set of motivational items used here has revealed two important dimensions of motivation: altruism and egotism. The following will discuss the findings in relation to these two underlying dimensions before addressing the issue of other motivations and concluding with the implications of the findings for survey researchers.

4.4.1 Altruistic motivation

The two main elements of the altruistic dimension are contributing to research and the survey topic. The findings are consistent with previous studies of participation motivation which highlight the importance of contributing to research (Bosnjak & Batanic 2002; Tuten *et al.* 1999) and the survey topic (Bosnjak & Batanic 2002; Edwards *et al.* 2002; Evangelista *et al.* 1999; Goyder 1987; Heberlein & Baumgartner 1978; Martin 1994; McDaniel *et al.* 1987; Van Kenhove *et al.* 2002).

The opportunity to contribute to research was the most important motivational factor, cited by three-quarters of respondents. When men were clustered into groups according to how they rated the motivational items, this item remained important for members of all three clusters, including men who had a tendency to consider other items as less important. The cluster analysis also showed that while respondents were motivated by both altruistic and egotistical considerations, altruistic motivation was more important across all three cluster groups. Although older, gay and HIV tested men were somewhat more likely to have higher scores on the altruistic dimension, these variables accounted for very little of the variation in respondents' scores of altruism.

The second most important motivation was the topic of the survey. The 0.45 correlation between an interest in the topic and contributing to research indicates that there is a degree of consistency between how men rated these two items. Moreover, the profile of the men for whom these items were important was similar. They were more likely to self-identify as gay, be more open about their sexuality, to have tested for HIV and to read the gay press. Knowing gay men with HIV also increased the

importance of contributing to research and a greater degree of socialising with gay men increased interest in the survey topic. These variables of gay identity and affiliation are inter-related, and it seems likely that the more gay-affiliated men consider both the topic of the survey and their contribution to research as relating to sexual health promotion and HIV prevention, which are particularly important issues for gay men. The principle of social validation appears to be operating in conjunction with the reciprocation principle and the topic of the survey in encouraging participation, supporting the argument that the interaction between the characteristics of the sample and the impact of the topic are crucial to response (Goyder 1987).

The decision to complete the motivational section of the survey was also associated with the gay-affiliation profile described above. Men who completed the motivational section were more likely to be gay, open about their sexuality, have sex with men only, know gay men with HIV, be members of gay groups, read the gay press, visit gay venues and socialise with gay men. The logistic regression supported the importance of sexual behaviour and identity, and indicated that age, education, taking longer to complete the main survey, being depressed and living outside London were also predictors of completing the motivational section. The fact that the men who went on to complete the motivational section tended to take longer in completing the main survey is an indication of their engagement with the Internet and HIV study.

There was some suggestion that the topic of the survey may not always be understood in the context of sexual health research. Whereas the older, gay affiliated and healthier men found the survey topic more motivating, men who had used online porn were less likely to consider this to be a motivating factor. Although this finding may be spurious, it would be interesting to explore whether the use of online porn colours the perception of a survey about "Sex, Health, Internet", such that it is less likely to be seen as a motivating factor.

The third element of the altruistic motivational dimension was the importance of the support of the university, which was important to less than half of the men and correlated only moderately with contributing to research and the survey topic. As with the other two elements of the dimension, men who rated it as important were more likely to be gay and to have gay affiliations. It is no surprise that a higher social class and higher education were associated with a belief in the importance of university support, but the reasons behind the range of other associations are less obvious and not pursued here. The importance of invoking the principle of authority should not be underestimated. It was found to be an important factor in a systematic review a number of studies (Edwards *et al.* 2002) and Goyder's (1987) work has shown how government sponsorship led to a marked increment in response, even though only one fifth of respondents to his survey on surveys felt that government sponsorship would be important to their response.

Respondents who volunteered their own motivations for participating were also more likely to cite altruistic reasons and contributing to research as important motivational factors. The data indicate that there was strong consensus among the respondents on the importance of contributing to research, which is consistent with Groves and Couper's (1998) argument that the reciprocity rule is likely to be of key importance to survey participation. The findings are also consistent with a measure of independence between social role and attitude towards survey participation (Goyder 1987).

4.4.2 Egotistical motivation

Liking filling in questionnaires loaded most highly onto the egotistical dimension and was the least persuasive of all motivational items, which less than one fifth of men found important. It correlated particularly poorly with an interest in the survey topic and was only weakly related to university sponsorship and contributing to research. This reinforces the idea that the purpose of the survey played a key role in triggering the desire to participate – men for whom this was important were more likely to be discriminating about which surveys they complete and less likely to

enjoy filling in questionnaires in general. The subgroup differences do not add any further to the interpretation of the data.

The respondents also thought that the look of the pop-up or banner was generally of minor importance in their decision to participate, with just one quarter of men saying it was important. Again, subgroup differences do not bear further discussion. The finding indicates that the principle of liking may lose its persuasive power in the absence of a person as its object and concurs with previous findings that certain elements of design are not of key importance to survey compliance (Briggs & Hollis 1997; Dillman *et al.* 1998; Helgesen *et al.* 2002). It would be interesting to test this finding through manipulation of the appearance of the pop-up or banner.

Explicit reference to the sexual content of the survey produces a different response from an interest in the general topic of survey. Although the topic may induce an altruistic desire to contribute to research, the act of filling in a survey of sexual behaviour is also ego-driven. One of the participants in the qualitative interviews[2] alluded to this dual quality when he said, *"you know, it's like a Cosmopolitan questionnaire but with more cred"*. The opportunity to complete a survey of sexual behaviour was important for nearly half of the men and significantly more important for men who did not hold a degree, had sought cybersex, were in better health, lonely and spent more time online cruising. The relationship between sexual experience and filling in surveys of sexual behaviour will be examined in Chapter 6. It has been found in a number studies (Bogaert 1996; Catania *et al.* 1990; Dunne *et al.* 1997; Strassberg & Lowe 1995) and bears further investigation.

The opportunity to complete an online survey was related to liking the opportunity to complete a survey of sexual behaviour and was important for about half of the men. It was also related to other online behaviours. Men who enjoyed this opportunity were more likely to have sought cybersex, completed the 2002 web survey, finished the survey speedily and spend longer on the Internet per week. There was no significant association between speed of survey completion and amount of time spent

[2] See Chapter 8 for further details of this piece of analysis.

on the Internet per week, indicating that some men appreciated the opportunity to complete the web survey which they were able to do quickly despite, or more likely because of, the greater amount of time that they spent on the Internet per week. The relationship between time spent on the Internet and time spent completing the survey merits further investigation. As Lukawetz (2002) has argued, more time spent online increases the likelihood of exposure to surveys, allows more time to complete them and may lead to stronger identification as an Internet user, encouraging web survey participation.

Only half of the men had taken part in other online surveys, but those who had done so were more likely to believe that the opportunity to take part in an online survey was a motivating factor than men who had not and had participated in an average of 2 such surveys. Although the novelty effect cannot be discounted, a German study found that the average person participated in less than one survey per year (Stocké & Langfeldt 2004), indicating that these UK men may be more highly involved in online research and that the principle of consistency is influencing their compliance.

The support of gaydar and gay.com was important to less than half of the men. There were a number of socio-economic differences associated with finding the support of gaydar and gay.com more or less persuasive and, interestingly, men who came from gaydar found the support of these websites more important than men who came from gay.com. Although knowing gay men who were HIV positive tended to reduce the importance of this item, it tended to be increased for men who had greater affiliation with the gay community, both online and offline. Thus, its importance was associated with reading the gay press and visiting gay venues, as well as with more frequent and lengthier online cruising. This is further evidence that the principle of social validation was invoked.

A number of variables significantly predicted respondent scores on the egotistical motivational dimensions, but accounted for little of the variation in these scores. The altruistic dimension has greater cohesion than the egotistical dimension, with items clearly related to an altruistic wish to contribute to important academic research.

Although the items that load onto the egotistical dimension are more diverse, the above summary has sought to draw out their similarities in terms of the egotistical motivation that underpins them.

4.4.3 Other aspects of motivation

The data relating to online survey participation exhibit a degree of consistency which adds validity to the findings. Men who had taken part in the 2002 Internet and HIV web survey were more likely to have taken part in other online surveys. Men who spent longer on the Internet per week and who had been using the Internet for longer to look for sexual partners were also more likely to have taken part in other online surveys. Men who had completed other online surveys were also more likely to have completed the survey in the first week which is evidence of a readiness to participate among these men and is consistent with Nederhof's (1986) finding that respondents with more research experience responded with less hesitation.

The data also suggest that respondents may have completed online sex surveys before. As described above, previous online survey completion was associated with taking part in the 2002 Internet and HIV web survey and with using sites like gaydar and gay.com, which field in-house customer satisfaction surveys, for longer. Men were also more likely to have used online porn, where "pop" surveys of sexual behaviour may feature. It would be interesting to explore whether respondents have taken part in other sex research in order to examine the extent to which the principles of consistency and scarcity apply to this activity.

Men were extremely unlikely to know anyone else who had completed the survey, suggesting that was not a good indicator of social validation. The item provided further evidence, however, of the internal consistency of respondent reporting. An openness about sexuality and greater exposure to other gay men, through having a male partner, knowing gay men with HIV, socialising more with gay men and visiting gay venues, resulted in a greater chance of knowing someone who had completed the survey. These men were also more likely to have tested for HIV and to read the gay press, which are indicators of a greater gay affiliation.

Men who had participated in the 2002 Internet and HIV web survey were more likely to know other men who had taken part, suggesting that the survey or similar surveys may have been discussed prior to participation. Such discussion may have been online, as men who logged on to cruise more often had a greater tendency to know other men who had taken part. Men who completed the survey in the small hours were unlikely to know anyone else who had participated.

Findings from the question about seeing the pop-up or banner before were validated against both external data and self-reported behaviour. Thus, men who had seen the pop-up or banner before were more likely to take part in later weeks. They were also likely to report that they had been logging on to sites such as gaydar and gay.com over a longer period and that they did so more frequently and for longer.

There was an indication that some men may have been committed participators who did not necessarily take part the first time that they see the pop-up or banner, but waited for an appropriate time. Men who did not participate the first time they saw the pop-up or banner were more likely to take part in the morning and to spend longer filling in the main survey. They were also more likely to have taken part in the 2002 survey and were perhaps waiting for the right opportunity to repeat their contribution. However, there was also evidence to suggest that some of the respondents did not interpret the question about seeing the pop-up or banner before as referring specifically to the invitation to complete the survey that they had just finished. They may have thought it referred also to 2002 Internet and HIV web survey or invitations to participate in other online surveys of sexual behaviour.

In conclusion, the findings show that it is helpful to conceptualise participation motivation in terms of its altruistic and egotistical dimensions. Although elements of the compliance model were found to encourage participation, an exploration of the model's validity in relation to participation motivation could be usefully undertaken using psychometric principles.

4.4.4 Implications for researchers

The findings suggest that participation was driven by respondent altruism. The small magnitude of the differences between subgroups in the bivariate analyses and the low level of variance explained by independent variables in the regression analyses show that contributing to research and the survey topic were important motivating forces for the majority of the sample, regardless of individual characteristics. Emphasising the opportunity to contribute to research into sexual health and HIV prevention may help to increase participation in web surveys of sexual behaviour among gay and bisexual men, but such appeals should be carefully worded so as not to alienate men who have less affiliation with the gay community.

5 Web survey drop out

Synopsis

Drop out from the 2003 Internet and HIV web survey was examined through secondary analysis of drop out patterns and analysis of primary data on respondent reasons for quitting. Most of the drop out occurred in the early stages of the survey and the analysis identified minor differences on key variables between those who completed and dropped out. Time available was implicated as an important cause of drop out.

5.1 Introduction

This chapter turns from why respondents *began* the 2003 Internet and HIV web survey to what may have caused them to *quit*, with the aim of understanding how such behaviour comes about and the possible biases it may introduce into the sample (Ross *et al.* 2004). It addresses the question:

"Among gay and bisexual men, who drops out of web surveys of sexual behaviour, at what point and why?"

Self-completion is a favoured method of collecting data on sexual behaviour, and yet we know little of what happens after such questionnaires are distributed. Indeed, we know very little about how respondents receive or react to self-completed surveys in general and, as Bosnjak and Bandilla (2000) have argued,

> *"In traditional mail surveys, the response process basically remains a mystery. We do not know whether a potential respondent received the questionnaire at all, read it, and began answering it"* (Bosnjak & Bandilla 2000: 5).

Helgeson (1994) sought to address this gap in the research through a phenomenological exploration of what happens when respondents receive postal surveys. Two weeks after receiving a market research survey on alcohol, respondents reported what had happened when they received the questionnaire. The

findings indicated that respondents either threw out the survey or filled it in immediately, left it and eventually lost it, or filled it in and failed to send it back. The small scale of the survey and the approach taken meant that the incidence of these behaviours was not quantified and questions remain about the self-completion process.

The advent of the Internet has, however, has made it possible to open up what Bosnjak and Bandilla term the "*black box*" of survey response and explore what is going on in much greater detail. What distinguishes web surveys from traditional pen-and-paper self-completion surveys is that they can routinely collect paradata on survey behaviour that was not previously easily captured. We can, for example, record the exact time that the respondent started and finished the survey and, as this chapter describes, we can identify the point where non-completers quit.

Whereas survey non-completion has been typically associated with telephone surveys (Kalton 1983), the use of web surveys allows us to extend our knowledge of how and when it happens in a self-completion mode. Although I would not wish to generalise the findings to other modes - the facility to preview questions in a pen-and-paper survey, for example, is likely to filter out respondents who are less interested in the survey topic from the start (Dillman 1978) – the findings may provide some insight into a process that has essentially remained beyond research.

This introduction begins by describing the ways in which survey response may be classified, with particular emphasis on a recent typology of web survey response (Bosnjak & Bandilla 2000). It goes on to describe the empirical evidence on factors that have been found to influence dropping out of a web survey.

5.1.1 Classifying survey response

Three main types of survey response based on traditional survey techniques have been identified – unit nonresponse, partial or item nonresponse and complete response (Bosnjak & Bandilla 2000). Unit nonresponse occurs when data are not collected from an individual (or other sampling unit) that was included in the

sampling frame and partial or item nonresponse refers to responses containing missing data.

Bosnjak and Bandilla (2000) took advantage of the paradata capturing properties of the Internet to devise a new typology which describes seven types of responders to web surveys. The typology is based on analysis of the log-files and data gathered from 1,469 respondents who were invited to participate in a survey through search engines such as Yahoo and Altavista. They were able to reconstruct responses to the survey by using a programme which allowed respondents to skip questions and by adopting a screen-by-screen design where each page of the questionnaire is downloaded separately.

The following typology was identified and is illustrated in Figure 5.1 (on the following page). The prevalence of each response behaviour is inserted in brackets after the each category:

i) **Complete responders (25.3%)**
view and answer all questions

ii) **Item nonresponders (36.0%)**
view all questions but do not answer all of them

iii) **Lurkers (6.9%)**
view all questions but do not answer any of them

iv) **Answering dropouts (4.3%)**
answer the questions that they view and then quit

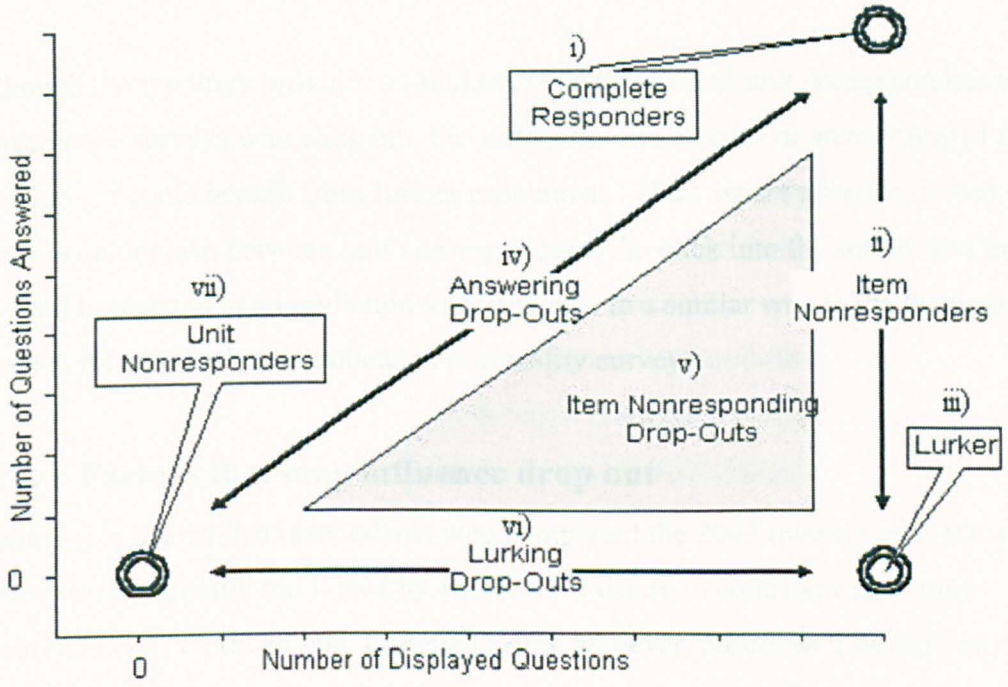
v) **Item nonresponding dropouts (4.0%)**
view some of the questions, answer some of them and then quit

vi) **Lurking dropouts (13.3%)**
view some of the questions and then quit

vii) **Unit nonresponders (10.2%)**
do not answer any of the questions

As the above data show, complete responders and item non-responders make up 61.3% of those who entered the survey and represent those who finished it. The responding drop out group made up 8.3% of potential respondents and the remaining 30.4% provided no responses to the survey.

Figure 5.1: Types of response in web surveys and their prevalence



Source: Bosnjak and Bandilla (2000)

Figure 5.1 illustrates the continua between unit nonresponding, complete responding and lurking, and suggests that,

> *"Answering drop-outs may be easily converted into Complete responders if we develop an understanding of the reasons behind the choice to end participation"* (Bosnjak & Tuten 2001: 9).

It is with this in mind that the following analysis of drop out from the Internet and HIV web surveys was undertaken.

The typology can also be used to illustrate the distinction between volitional and non-volitional drop out (Bosnjak 2001) whereby dropping out is the result of volitional withdrawal or factors beyond the respondents' control, such as technical difficulties. Thus, it seems likely that respondents who reached the end of the survey did so at their own volition, including (i) complete responders, (ii) item nonresponders and (iii) lurkers, whereas non-completion may have been beyond the control of (iv) answering dropouts, (v) item nonresponding dropouts, (vi) lurking dropouts and (vii) and unit nonresponders.

Although the typology provides a satisfactory description of unit nonresponders to convenience surveys who click into the survey but fail to view or answer any of the questions, it could benefit from further refinement. Thus, where possible, it may be better to distinguish between unit nonresponders who click into the survey and those who fail to respond to an invitation to participate, in a similar way to the distinction between refusals and non-contacts in probability surveys.

5.1.2 Factors that may influence drop out

Chapter 4 indicated that respondents who completed the 2003 Internet and HIV web survey were originally motivated by an altruistic desire to contribute academic research into HIV prevention. Catania (1997), however, describes how motivation is likely to change over the course of an interview with the respondents' growing awareness of the survey topic. What follows is an examination of the empirical evidence on factors that have been found to change motivation or otherwise influence respondents to the extent that they then drop out of the survey.

Such factors may be divided into those that are survey-driven by, for example, the nature of the questions or the length of the questionnaire and those that are situational (Knapp & Heidingsfelder 2001) where, for example, the respondent is interrupted by a third party. The survey-driven features may be manipulated whereas situational features are beyond the survey designer's control. An interaction between survey-driven and situational factors may occur where, for example, a questionnaire

contains features that take longer to download and the respondent's Internet connection is slow.

Although a complete explanation of survey drop out would include both the survey-driven and situational features, the impact of the latter is not easy to quantify. Analysis of data from web surveys has done a great deal to open up the "black box" of survey response but the impact of situational factors remains difficult to capture. Many of the following studies, therefore, have sought to explore response behaviour through the experimental manipulation of variables that are within the researcher's control and few have sought to explore the situational context. Reips (2002a) acknowledges the importance of situational factors but argues that drop out is only problematic if it introduces bias – if, for example, people drop out because they don't like the topic of the survey. Although the apparently random nature of situational factors *may* not affect the validity of the findings, it would be unsafe to assume that such factors operate independently and do not, therefore, contribute to bias. A respondent who is highly motivated by the topic may, for example, be less influenced by situational factors which may distract a less motivated respondent.

5.1.2.1 Amount and point of drop out

The amount of drop out from web surveys has been found to vary considerably. This may be partially attributable to a lack of clarity over how to calculate drop out rates, as implied by the earlier discussion of unit nonresponders. A survey of investigators who had conducted twenty online experiments revealed completion rates of between 13% and 99% for their online questionnaires, with a mean completion of 66% (Musch & Reips 2000). Similarly, 71% of students who responded to an email invitation to participate in a web survey returned a complete response (Crawford *et al.* 2001) as did 66% of Latino MSM (Ross *et al.* 2004) and 49% of Swedish men (Ross *et al.* 2003) who responded to the advertisements for two separate web surveys of sexual behaviour with similar designs to the Internet and HIV survey.

The greatest drop out typically occurs in the early stages of the survey (Jeavons 1998; Knapp & Heidingsfelder 2001; Lindley *et al.* 2003; Ross *et al.* 2003; 2004).

This is consistent with the foot-in-the-door phenomenon (Freedman & Fraser 1966) whereby people are more likely to complete a more demanding task if they have already completed a similar but less demanding one. Thus, respondents who had previously participated in a brief telephone survey were significantly more likely to participate in a second longer survey than those who had not previously been approached (DeJong 1981) and respondents who have completed a few questions will be more likely to complete the entire survey.

Response behaviour was found to be remarkably similar for web surveys on three different topics, with the first question likely to cause refusal or confusion in all cases (Jeavons 1998), and an exploration of drop out in nine web surveys suggested that half of all drop out is likely to occur at the first question, with the first two questions resulting in the highest drop out (Knapp & Heidingsfelder 2001). Dropping out of a web survey of sexual behaviour was also described as following a curve of negative acceleration, with high drop out occurring early in the questionnaire (Ross *et al.* 2003). The evidence clearly indicates that the highest drop out from the Internet and HIV web surveys is likely to occur during the first few questions.

5.1.2.2 Commitment

Studies manipulating the effort required by respondents to enter the survey throw some light on why so many respondents drop out in the early stages. The surveys upon which the above studies were based appear to have been accessed with the click of a mouse. Respondents who had to manually enter a PIN code in order to participate in a survey were more likely to complete a greater number of blocks of survey items than those who logged in automatically (Heerwegh & Loosveldt 2002b). In the manual log-in condition, respondents dropped out in a steadier linear pattern, whereas those in the automatic condition were more likely to drop out in the earlier sections of the survey than the later sections. Manual log-in thereby appears to filter out the less committed respondents.

Similarly, students who had to log in manually to a web survey were less likely to begin the survey than those who were logged in automatically, but having got over

this hurdle, respondents in the manual condition spent more time filling in the survey (Crawford *et al.* 2001). This indicates that manual log in may be associated with more careful consideration of questions, potentially leading to higher data quality. Automatic log in may encourage some respondents to peruse the first few questions and drop out if they are not interested or continue to fill in the questionnaire but with a lower level of engagement. In relation to this, Knapp and Heidingsfelder (2001) found evidence that drop outs were also more likely to satisfice – they were less likely to complete open-ended questions and were quicker in answering them than those who completed the survey.

Where information is provided on survey completion time, the anticipated burden of completing a longer survey has an effect that corresponds to the actual burden of logging in manually. Crawford *et al.* (2001) obtained a lower response rate for those who were told that the survey would take longer to complete (20 minutes rather than 8 to 10 minutes) but the 20-minute group were less likely to drop out once they had started, resulting in a similar rate of completion across treatments. The length of time before drop out was similar for the two groups. The manipulation appeared to have a stronger effect on likelihood of response than on likelihood of drop out, indicating that other factors come into play once respondents have begun to fill in the survey.

The above studies have shown that reduced drop out appears to be symptomatic of greater commitment to the survey and that engaging such commitment may help reduce drop out. This is not easily done, however, and an attempt to do so by asking respondents to hit a button saying "Yes – I want to do the experiment" had no effect on drop out (Frick *et al.* 2004). The findings suggest, however, that respondents who were more committed to the Internet and HIV study would be less likely to drop out.

5.1.2.3 Time and the perception of burden

Crawford *et al.* (2001) were surprised to find that inclusion of a progress indicator increased drop out. They considered this to be an artefact of the survey. Respondents were required to answer a number of burdensome open questions

towards the beginning of the survey, at which point the progress indicator showed that they still had a large proportion of the survey left to answer, without indicating that the later questions would take less time. A subsequent study by Conrad *et al.* (2003) showed how sensitive respondents may be to their assessment of progression through the survey. Manipulation of the progress indicator led respondents to perceive that their early progress was fast (encouraging) or slow (discouraging) in relation to actual progress, or undisclosed. Encouraging news given at the beginning of the survey was found to decrease drop out and improve respondent perceptions of the survey.

This study indicates that respondents who have started a survey remain susceptible to the anticipated burden of longer completion time. Ross *et al.* (2004) argue that attention span may be a particularly important factor in web survey completion, given that more interesting sites are just a click away and the following studies provide further evidence of how completion time is implicated in drop out. Thus, business surveys that were lengthier or took longer to complete resulted in lower completion rates (MacElroy 2000).

In the same way, the time taken to download the web survey has also been implicated in drop out. Although unaffected by the respondent's web browser or operating system (O'Neil 2002), a slower Internet connection has been associated with greater drop out (Ross *et al.* 2003). An experimental investigation of survey design effects by Couper *et al.* (2001) showed that the inclusion of a progress indicator resulted in a 90% completion rate compared to an 86% completion rate where no progress bar was included. This marginal and non-significant effect led the authors to further investigation which showed that respondents in the progress indicator condition took significantly longer to complete the survey. This was attributed to increased download time which was thought to have mitigated the motivating effect of the progress indicator. Drop out is similarly likely to have occurred due to the increased download time of radio buttons (Heerwegh & Loosveldt 2002a) and logotypes (Lozar Manfreda *et al.* 2002a)

Dillman *et al.* (1998) examined the difference in respondent behaviour towards "plain" and "fancy" web surveys, with the aim of testing the use of conventional questionnaire designs. The "fancy" design included a number of features such as bright colours and dynamic formatting with result that it appeared less conventional but it also took longer to download. The findings showed that the plain questionnaire was more likely to be completed (93% vs 82%) and that its respondents were more likely to complete a larger number of questions before quitting (166 vs 156) and provide more answers to open-ended questions. Although the study appeared to support the use of conventional designs, it has been pointed out that the greater time needed to download the "fancy" survey is confounded with its unconventional appearance.

An extensive literature on question design in traditional surveys (eg Dillman 2000; Fowler 1993) has highlighted the importance of ensuring that questions do not demand too much time or effort. It is therefore no surprise that complex questions have been found to lead to survey drop out (Jeavons 1998) and that open-ended or multiple-choice formats result in a greater likelihood of quitting (Knapp & Heidingsfelder 2001). The indication that question burden increases drop out is further illustrated by the finding that the questions which had the highest drop out rates took longer than average to complete (Knapp & Heidingsfelder 2001).

The above evidence indicates that respondents may become impatient and quit the survey if they feel it is taking too long but this may not be such a problem where more time is available. Thus, respondents who participated in a web survey over the weekend were less likely drop out than those who participated during the week, particularly at the beginning of the survey (O'Neil 2002). The issue of time available is clearly an important question which the following analysis of drop out from the Internet and HIV web surveys will explore.

5.1.2.4 Personal information

Another factor that may influence drop out is the nature and location of questions asking for personal information. Questions about demographic details, particularly

requests for email identification, may lead respondents to quit the survey (Jeavons 1998) and drop out can be reduced by making such questions optional (O'Neil 2002).

It is, however, advisable to ask such questions at the beginning rather than the end of the survey, as this is associated with decreased drop out (Musch & Reips 2000; O'Neil 2002) without affecting responses to questions that may be influenced by social desirability bias (Musch & Reips 2000). The success of locating personal questions at the beginning of the survey has been viewed in the context of the common practice of asking users to register prior to access to websites (Bosnjak & Tuten 2001).

Whereas requests for personal information have been shown to have an important influence on drop out, the use of personal information in the invitation to participate in the survey does not have a similar effect. Thus, a personalised salutation was found to increase response to a web survey but was not associated with likelihood of drop out (Heerwegh *et al.* 2004).

Although the impact of specific questions of a sexual nature on web survey drop out remains largely unexplored, a telephone survey of sexual behaviour indicated that drop out was unrelated to individual questions (Catania *et al.* 1996). This suggests that particular questions of sexual behaviour in the Internet and HIV web survey are unlikely to provoke drop out.

5.1.2.5 Use of incentives

Although the Internet and HIV study did not use any material incentives, there is a growing literature on the use of such incentives in web surveys, as described in the Chapter 4, which indicates that the optimal online incentive has yet to be devised. The use of such incentives has, however, been found to reduce the likelihood of drop out in a number of studies (Frick *et al.* 2001; Göritz 2004 cited in Göritz 2004; MacElroy 2000; Musch & Reips 2000; O'Neil 2002). Where respondents received some kind of financial incentive (monetary or lottery) completion rates were higher than where no rewards were offered (Musch & Reips 2000) and even where lottery

incentives did not affect motivation to start the survey, they decreased the percentage of respondents who dropped out (Frick *et al.* 2001).

A meta-analysis by Göritz (2004 cited in Göritz 2004) indicated that increasing the value of incentives reduced drop out for both panel and one-off web surveys. Her own experiments did not support this finding (Göritz 2004), however, and increasing the $5 cash incentive to at least $22 in web-based business-to-business surveys was also found to have little further effect (MacElroy 2000). Bosnjak and Tuten (2003) found that respondents offered a prize were the most likely to both enter and least likely to quit a web survey, followed by the group that was promised payment upon completion, then those who were offered payment at the outset (via PayPal) and finally the control group. Although Ross *et al.* (2004) also found an association between the incentive and completion rate, their respondents chose their own incentive making it impossible to disentangle cause and effect.

A different incentive, course credit, has also been shown to encourage survey completion (O'Neil 2002). Having clicked into the study, 67% of students participating for course credit completed it, compared to 19% of members of the general public who accessed the study via four psychological research websites. Of course, particular characteristics of these students may have also increased their likelihood of survey completion in comparison to the other group.

Although comparison of the characteristics of respondents who complete and drop out of web surveys would seem to be of value, the research has tended to concentrate on the survey-driven factors that are implicated in drop out, as described above, or to use drop out as an indicator of respondent burden in relation to features of survey design (Coates 2004; Conrad *et al.* 2003; El-Menouar & Blasius 2004; Kaczmirek *et al.* 2004). Of particular relevance to the present study, however, is the small number of studies that have compared the characteristics of completers and drop outs in web surveys of sexual behaviour, to which I now turn.

5.1.2.6 Respondent characteristics

Whereas Lindley *et al.* (2003) found no significant differences in the demographic profile of the LGBT college students who quit or completed their web survey of sexual behaviour in the US, Ross *et al.* (2003) found a number of differences in their examination of drop out from a Swedish survey of sexual behaviour. Men were more likely to drop out than women and did so significantly earlier in the questionnaire. Thus, half of the men who began the survey dropped out compared to 43% of the women. Half of the male drop outs had quit by item 25 whereas half of the female drop outs reached item 49. The data also suggested that gay and bisexual men were less likely to drop out than those who were heterosexual, as were men who were younger, better educated, urban-dwelling and living alone.

A subsequent study of Internet sex among Latino MSM in the US (Ross *et al.* 2004) highlighted the importance of the survey topic, with greater drop out found among men for whom the topic was likely to be less relevant. Thus, men were more likely to drop out if they self-identified as bisexual or heterosexual rather than gay, had not had sex with a man they met on the Internet or were in a seroconcordant monogamous relationship. This study found no significant effect of age. It concluded with a recommendation that the findings should be further tested on different populations, which the following analysis will undertake through its investigation of drop out among gay and bisexual men in a UK web survey of sexual behaviour.

5.1.2.7 Internet behaviour

Dropping out of a web survey may also be connected to other aspects of the respondent's Internet behaviour. Survey drop out has been found to increase as the length of time in the field increases, with three surveys covering different topics resulting in an 11% drop out over a 7-day fielding, compared to 24% over 19 days and 30% over 30 days (Jeavons 1998). A survey that is fielded for a shorter time is likely to reach a higher proportion of heavier Internet users whom Lukawetz (2002) has argued will have more occasion and stronger motivation to complete the survey. It is also more likely to reach the regular users of the site to which it is attached who

are more likely to be "fans" of the site and to complete the survey (Knapp & Heidingsfelder 2001) and the importance of this relationship between the respondent, the site and the survey may explain why Ross *et al.* (2003) did not find an association between greater use of the Internet in general and drop out.

In summary, the literature shows that drop out is most likely to occur in the early stages of the survey. It is prompted by a range of factors and particularly likely to occur if respondents find the response process overly time-consuming. Although the researcher may increase completion through careful design and the provision of incentives, the relevance of the topic is likely to be of key importance. Much remains to be learnt about the factors that influence drop out and little attention has been paid to the impact of respondent characteristics, which the following seeks to address in the context of drop out from the Internet and HIV web survey.

5.2 Methods

This section begins with a discussion of the issues surrounding identification of the drop out point in the Internet and HIV 2003 web survey. It goes on to describe the drop out comparisons that were made on the basis of subgroup membership, followed by a detailed description of the data analysis techniques used in these comparisons. It concludes with a description of the drop out pop-up and items included in the motivational section described in Chapter 4 that provide further data on drop out.

5.2.1 Drop out point

The 2003 web survey was designed so that unless questions needed to be answered for routing purposes, respondents could skip a question and proceed to the next one. Additional paradata were collected with the individual's responses to the questionnaire, including the time that the survey was started and completed. If the respondent dropped out before reaching the final page, his final time could not be recorded. Therefore any respondent whose final time was not logged is classified as having dropped out.

The dataset allows us to identify the last question that the respondent answered but it does not allow us to pin point the last question that the respondent may have viewed. If the question following the respondent's last question was compulsory, respondents may have done one of the following:

- dropped out without proceeding to the following question;
- proceeded to the following question, didn't answer it and dropped out;
- proceeded to the following question, tried to skip it and then dropped out.

If the question following the last question answered was non-compulsory, respondents may have done one of the following:

- dropped out without proceeding to the following question;
- proceeded to the following question, didn't answer it and dropped out;

- proceeded to the following question, skipped it and proceeded further before dropping out at a non-compulsory question;
- proceeded to the following question, skipped it and proceeded further before dropping out at a compulsory question.

Respondents cannot definitively be classified as having "dropped out" until they fail to answer a compulsory question – up to this point they could have been browsing questions or "lurking". For the purposes of the following analysis, however, respondents are deemed to have "dropped out" at the point where they fail to answer any further questions.

Men who had not been sexually active with men in the previous year were ineligible for the study and were therefore filtered out of the questionnaire. In order to filter out such men, the following questions were included between the sociodemographic questions and the main body of the survey, at questions 20 and 21:

- In the past year have you had sex with: men only; mostly men / occasionally women; men and women; mostly women / occasionally men; women only
- Are you sure you have not had any sexual contact with a man in the previous 12 months?

Although these men left the survey before the final question they cannot be considered to have dropped out and the following begins with a description of how the analysis takes this into account.

5.2.2 Comparison of subgroups

The inclusion of these filter questions raises the issue of whether the population of men represented by the sample responding *prior* to these questions is comparable to the population of men represented by the sample responding *after* these questions. It is possible that some of the men who dropped out before the filter questions had not been sexually active with men in the last year, whereas all of the men who dropped

out after the filter questions had been sexually active with men in the recent past. For this reason, the drop out was examined in the following way:

i) for the whole questionnaire;
ii) up to the point where the filter questions were asked (pre-filter);
iii) following the filter questions (post-filter).

Sociodemographic and sexuality questions were asked before the filter questions and questions about HIV status were asked immediately after the filter questions. Data from these questions were therefore more or less complete for men who dropped out after the filter questions and become increasingly complete for men who dropped out before the filter questions. Thus, the age of the respondent was the only data recorded for men who dropped out after answering the first question, age and ethnicity were the only data recorded for men who dropped out after answering the second question, and so on. This pattern continues in the following order, although more than one question may be asked for each of the variables listed and data were subject to item nonresponse for non-compulsory questions:

1st	age	7th	place of birth
2nd	ethnicity	8th	sexual orientation
3rd	work status	9th	openness
4th	education	10th	sex with only men
5th	place of residence	*11th*	*filter*
6th	type of area	12th	HIV status

Social class is not included because it requires manual coding and there were not sufficient resources to code the responses of all the men who completed these questions before dropping out. Questions about HIV status were asked just after the filter questions but this variable is included in the subgroup comparisons because it is a key variable in the analysis of the main Internet and HIV survey and its impact on drop out is therefore of empirical interest.

Although the analysis in the chapters which make up the remainder of Part II (Chapters 4 and 6) is based on UK samples, it was not possible to identify UK men until questions on residence had been asked over the course of the pre-filter questions. Thus, the following analysis of drop out for the full questionnaire includes all men who started the survey whether from inside or outside the UK, as does the analysis of the pre-filter questions. The analysis of drop out for the post-filter questions focuses on the UK sample in order to provide findings that can be compared with those from the other chapters.

5.2.3 Data analysis

Kaplan-Meier survival analysis (Kaplan & Meier 1958) was used to examine rates of drop out. Although survival analysis has many applications, including examination of duration of strikes or time taken to complete specified tasks (Cox & Oakes 1984), it is probably most commonly known for its capacity to explore the survival times of patients in clinical trials. Applying Kaplan-Meier survival analysis to explore drop out points from a web survey therefore takes a new approach that was first applied by Ross *et al.* (2003; 2004).

One of the features of the statistical technique that makes it suitable for analysing patient survival, is its ability to account for the fact that we cannot determine the future survival of patients who are still alive when the clinical trial comes to an end. Because excluding such patients from the analysis would bias the results, they are incorporated into a Kaplan-Meier survival analysis, through the technique of "right censoring".

In the analysis undertaken here, exclusion of respondents who completed the survey would similarly bias the results and such cases were therefore right censored in the following way:

i) **whole questionnaire**

men who completed the survey were censored;

ii) **pre-filter questions**

cases were censored that had not dropped out by the first filter question, on the basis that if the survey had terminated here, all respondents reaching this point would have been completers;

iii) **post-filter questions**

men who completed the survey were censored.

The Kaplan-Meier survival analysis has the additional advantage of not requiring that data are normally distributed which is often a characteristic of the data to which it is applied (Swinscow & Campbell 2002). The two functions which describe the survival distribution are the survival function and hazard function. Broadly speaking, the survival function represents the probability that the individual will survive from the starting point to beyond a future given point, and the hazard function is the risk of failure at a given point having survived up to that point (Collett 2003). Thus in terms of the present analysis, the survival function describes the probability that respondents will not drop out before reaching a particular question. Its plot is therefore equal to the observed drop out pattern and is used here as a visual representation of drop out.

The ***log rank statistic*** is used to compare the survival functions of different subgroups of respondents. In terms of the present analysis, it tests the null hypothesis that the observed drop out pattern for subgroups does not differ by comparing the observed number of drop outs at each question with the expected number of drop outs if the survey completion behaviour of the subgroups were the same (Landau & Everitt 2004).

Collett (2003) recommends use of the log rank statistic unless the estimated survival functions or previous data provide good reason to doubt the proportional hazards function. Violation of the assumption of proportionality is manifested when the plots of the survival functions cross. The Breslow statistic may be recommended for cases where the assumption of proportionality is violated (Collett 2003) but it was

generated for each of the following comparisons and found to be almost identical to the log rank statistic. It is therefore not included in the findings which follow.

Stepwise logistic regression, as described in Chapter 4, was used to identify a set of potential predictors for whether or not the respondent was likely to drop out following the filter questions. Variables were included in the model that were significantly associated with dropping out. Although Cox regression can be used to model survival data, it does not provide an indication of the amount of explained variance. Examination of the data suggested serious violation of the assumption of a linear relationship between continuous variables and the hazard function on which Cox regression is based and also indicated that use of the continuous variable "age" may violate the assumptions of logistic regression. The continuous variable provides the most accurate measure of age, however, and is therefore used in the logistic regression models presented below. It should be noted that recoding age into a dichotomous variable produced similar results and the predictive variables in the logistic regression model are identical to those indicated by Cox regression.

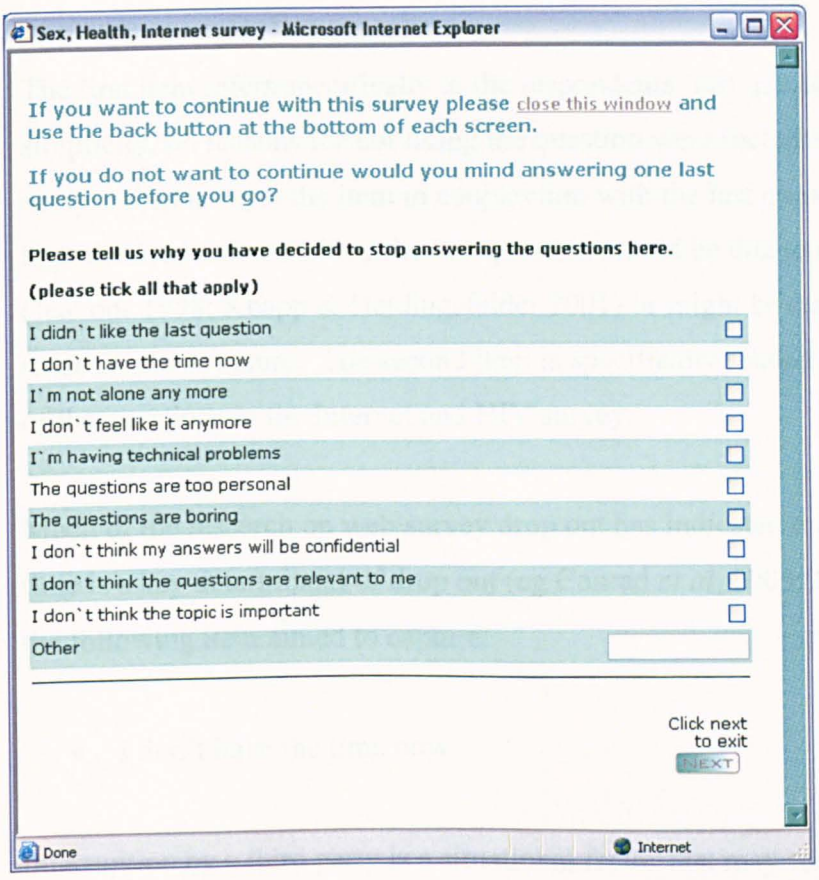
The ***chi-square test*** and ***t test for independent groups***, as described in Chapter 4, were used to examine whether there were significant associations between respondent characteristics and the likelihood of answering yes or no to seven routing questions. There was evidence that some of this analysis violated the assumptions of the t test but use of the Mann-Whitney U test produced the same results.

5.2.4 Drop out pop-up and drop out questions

5.2.4.1 Drop out pop-up

A number of extra features were incorporated into the 2003 Internet and HIV web survey. Paradata on date and time of completion and the website of origin (gaydar or gay.com) were recorded, the motivational section discussed in Chapter 4 was added and a drop out pop-up was included that appeared if men decided to quit the survey before they reached the end. If men closed down the survey, the pop-up (Figure 5.2 on the following page) appeared asking if they could complete one more question before leaving.

Figure 5.2: Drop out pop-up



A programming problem meant that the pop-up also appeared if men clicked on the back button on their toolbar instead of using the back button built into the survey – hence the instruction to close the window and use the back button at the bottom of each screen. Men were able to tick more than one box in response to the question.

Items in the drop out pop-up were drawn from the literature on web survey drop out, described in *5.1.2 Factors that may influence drop out* and survey compliance more generally in order to provide a comprehensive list of reasons for drop out. They were designed to be as simple as possible in order to minimise response burden and stimulate a response. The justification for including each item is described below.

The following two items relate to the design and nature of the questions:

- I didn't like the last question

- The questions are too personal

The first item refers specifically to the respondents' last question and, for the sake of simplicity, no reasons for not liking the question were included. The aim was to analyse responses to the item in conjunction with the last question answered. It was hypothesised that not liking the last question would be due to question design (Jeavons 1998; Knapp & Heidingsfelder 2001) or might be caused by questions of a more sensitive nature. The second item is specifically related to the sensitive content of the questions in the Internet and HIV survey.

Much of the research on web survey drop out has indicated that survey completion time is a key determinant of drop out (eg Conrad *et al.* 2003; MacElroy 2000) which the following item aimed to capture:

- I don't have the time now

Interruption by a third party is a situational factor that may cause the respondent to break off. The presence of other people during completion of the 1990 National Survey of Sexual Attitudes and Lifestyles was associated with an increased proportion of incomplete answers (Kupek 1998):

- I'm not alone any more

Attention span may be an important factor in drop out (Ross *et al.* 2004) and the following items were designed to capture a lack of engagement with the survey content. The Internet and HIV 2003 web survey is quite long, with a total of 158 questions, and some of the content is fairly repetitive in order to provide the level of detail required for analysis.

- I don't feel like it any more
- The questions are boring
- I don't think the questions are relevant to me

- I don't think the topic is important

Although every effort should be made to anticipate technical problems, servers crash and computer equipment becomes outdated as technology advances. Although having technical problems may preclude response to the drop out pop-up, the item was included in an effort to provide an thorough list of reasons for drop out.

- I'm having technical problems

The content of the Internet and HIV survey makes it particularly important that respondent anonymity and confidentiality should be ensured. Respondent concerns about privacy and confidentiality have been found to affect the likelihood returning the US Census (Singer *et al.* 1993) and it is hypothesised that an apparent lack of confidentiality may cause respondents to drop out.

- I don't think my answers will be confidential

Respondents were provided with the opportunity to express other reasons why they may have decided to quit at this point.

5.2.4.2 Drop out questions in motivational section

The motivational section described in Chapter 4 included the question "Apart from this time, have you seen the pop-up or banner inviting you to take part in this survey before?" Following this question, men who had seen the pop-up or banner before were asked how they had responded. If they did not take part they were asked why not and if they started the survey they were asked why they did not finish it. Items from these two questions were mapped onto the items in the drop out pop-up, so that there would be comparability across the "non-participation" type questions and if respondents had dropped out previously and filled in the drop out pop-up they would be able to use the same items, should they remember and wish to do so.

5.3 Findings

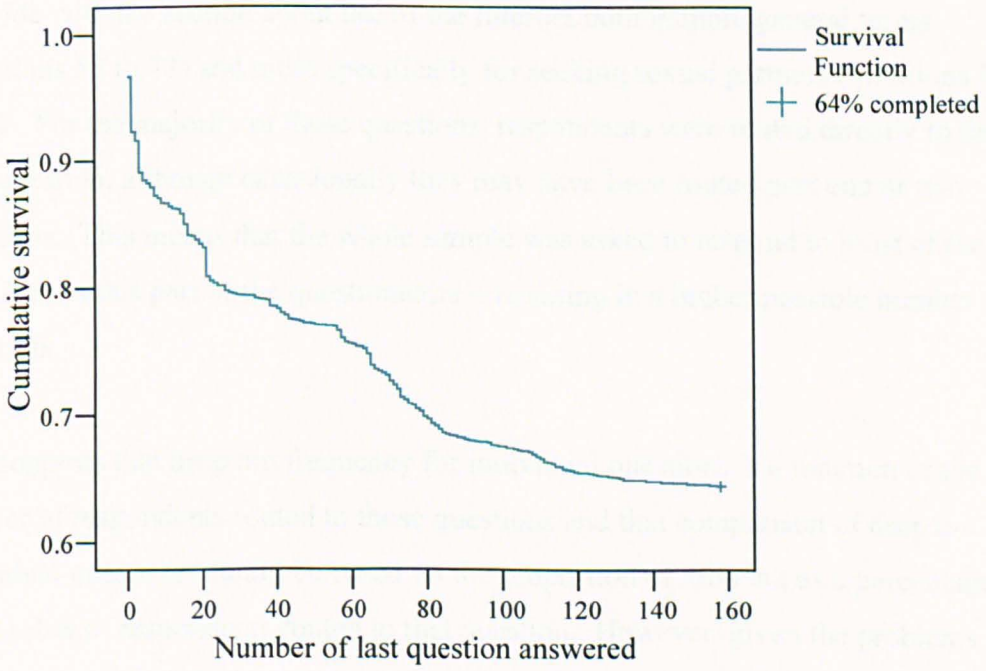
The findings from the secondary analysis of drop out from the Internet and HIV 2003 web survey will be presented first, followed by an analysis of the primary data on drop out that was collected with this survey.

5.3.1 Drop out over full questionnaire – all men

A total of 4,271 incidences of men starting the 2003 web survey were recorded, including responses from men who were not based in the UK. The survey was completed 2,752 times (64%) and in 1,420 cases (33%), men dropped out before the end. The remaining 99 men (2%) had not been sexually active with a man in the last year and were filtered out of the survey.

Figure 5.3 shows the observed drop out pattern for everyone who started the survey over the full length of the questionnaire.

Figure 5.3: Drop out 2003 over full questionnaire



The plot describes how men dropped out of the questionnaire from the starting point (at question 0) to the end point (at question 158) where 64% remained. It should be

noted that the cumulative survival for respondents whose last question was "0" (that is, those who did not answer a single question) is plotted at the mid-point between the 100% of respondents who clicked into the survey and the 93% who answered at least one question. The cumulative survival for the subsequent questions is plotted against the exact proportion of respondents remaining after that question. The same principle applies in the plots presented below.

The plot is characterised by a high rate of drop out during the pre-filter questions, numbers 1 to 19, with 83% of men remaining at question 19, dropping to 81% when men had been filtered out due to ineligibility. Half the total drop out occurred during these pre-filter questions, with the highest drop out occurring at the first question, which 289 (20% of drop outs) failed to answer. Following the filter questions at 20 and 21, the drop out rate slowed over the remainder of the questionnaire. The median last question for all those who dropped out of the survey was number 20.

Visual inspection of the plot also reveals an increased rate of drop out between questions 55 and 84. Examination of the questionnaire shows that these questions coincide with the section about use of the Internet both in more general terms (questions 55 to 72) and more specifically for seeking sexual partners (questions 73 to 84). For the majority of these questions, respondents were routed directly to the next question, although occasionally they may have been routed past one or two questions. This means that the whole sample was asked to respond to most of the questions in this part of the questionnaire – resulting in a higher possible number of drop outs.

This suggests that drop out frequency for individual questions is a function of the number of respondents routed to those questions and that comparison of drop out at individual questions should be based on the proportion of drop out as a percentage of the number of respondents routed to that question. However, given the problems described earlier with identifying an exact "drop out question" and because the preliminary analysis indicated that the amount of drop out for the majority of individual questions was very low, this analysis focuses mainly on comparing drop

out for different subgroups rather than different questions. The following will therefore examine patterns in the frequency of drop out over the course of the survey for different subgroups of respondents. It will compare drop out for the some of the subgroups identified in Chapter 4 over the pre- and post-filter sections of the survey. As described earlier, only subgroup classifications based on early questions or paradata could be used here.

There was a substantial amount of routing in the Internet and HIV web surveys and a possible limitation to this analysis is that it makes the assumption that the number of questions presented to individuals was randomly distributed across subgroups and not related to subgroup membership. The questionnaire included a progress bar which indicated the proportion of the questionnaire that the respondent had completed. If the number of questions was not randomly distributed, the higher drop out of certain subgroups might be prompted by the greater number of questions with which they were presented and the slower progress that they perceived. This issue is addressed below in *5.3.6 Influence of survey routing – UK men*, which suggests that higher drop out was not associated with the presentation of greater number of questions.

Table 5.1: Log rank comparison of drop out during pre-filter questions – all men

	Log rank statistic
Website of origin	13.6 **
Week of completion	27.3 **
Time of day	7.1
Age group	10.4 *
Ethnicity	15.5 **
Work status	1.4
Education	1.5
Place of residence	279.4 **
Place of birth	237.4 **

Note: *$p \leq 0.05$, **$p \leq 0.01$

5.3.2 Drop out over pre-filter questions – all men

This account of the comparative drop out of subgroups over the pre-filter questions for all men is followed by a similar description of the post-filter drop out which focuses on UK men.

Table 5.1 shows the results of applying the log rank statistic to examine whether there was a significant difference between the observed drop out patterns for the subgroups listed over the pre-filter questions. These comparisons do not include type of area, sexual orientation, openness, sex with only men or HIV status because the position of these questions within the survey meant there were not enough respondents upon which to base the analysis.

The data show how in the pre-filter part of the questionnaire, website of origin, week of completion, age, ethnicity and particularly place of residence and place of birth were significantly associated with likelihood of dropping out. All significant differences are described in detail below and shown in Figures 5.4 to 5.9 (on the following pages).

Men who entered the survey from gay.com were more likely to drop out during the pre-filter questions than men who came from gaydar, with 81.0% of men from gay.com finishing this section compared to 86.3% of men from gaydar. Men who completed the survey during the early weeks, particularly the first week, were less likely to drop out at the beginning of the questionnaire than men who completed it in later weeks. The completion rate for the pre-filter questions fell from 87.4% to 82.0% over the time that the survey was in the field. Older men were more likely to complete this section than younger men (92.2% vs 89.2%), as were white men compared to minority ethnic men (90.7% vs 83.1%), those living in the UK (97.6% vs 80.4%) and those who were born in the UK (98.8% vs 87.2%).

Figure 5.4: Pre-filter drop out 2003 by website of origin

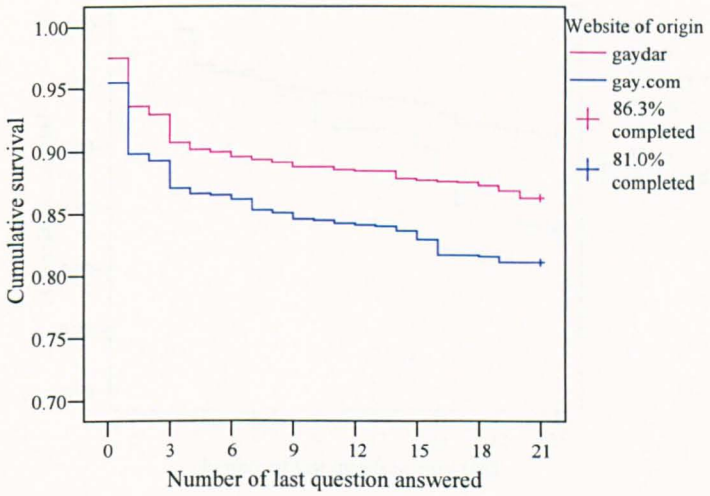


Figure 5.5: Pre-filter drop out 2003 by week of completion

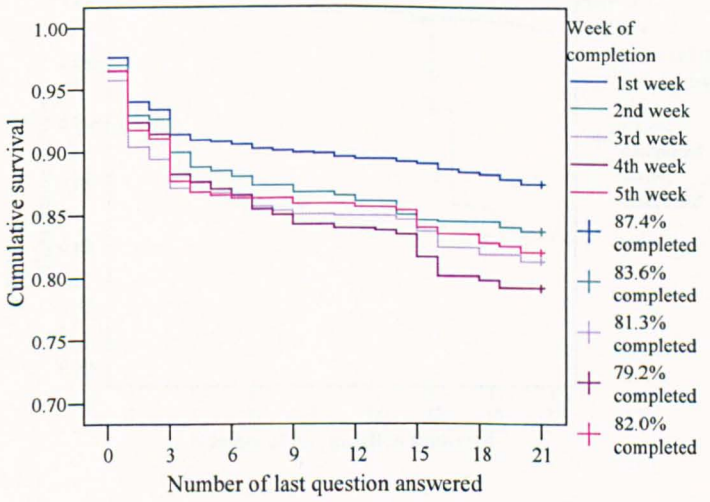


Figure 5.6: Pre-filter drop out 2003 by age group

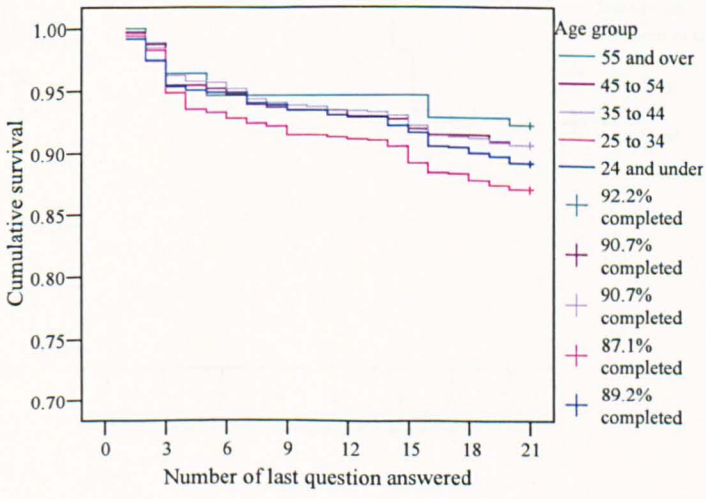


Figure 5.7: Pre-filter drop out 2003 by ethnicity

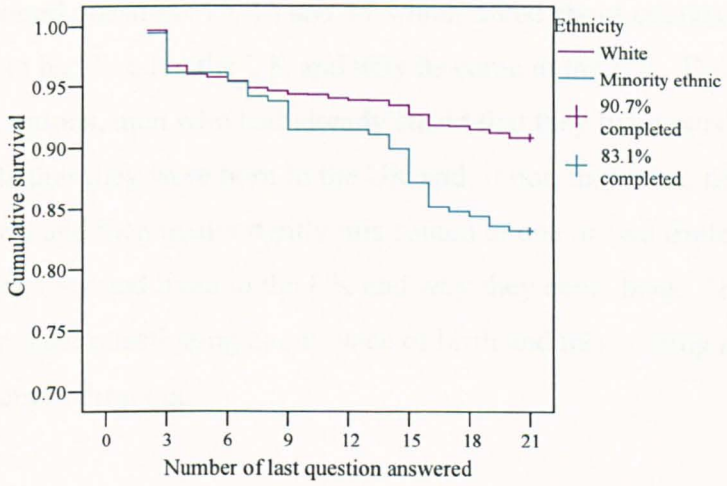


Figure 5.8: Pre-filter drop out 2003 by place of residence

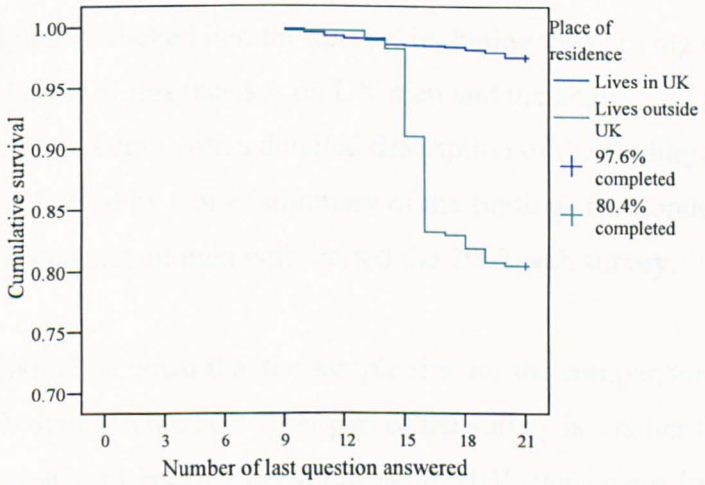
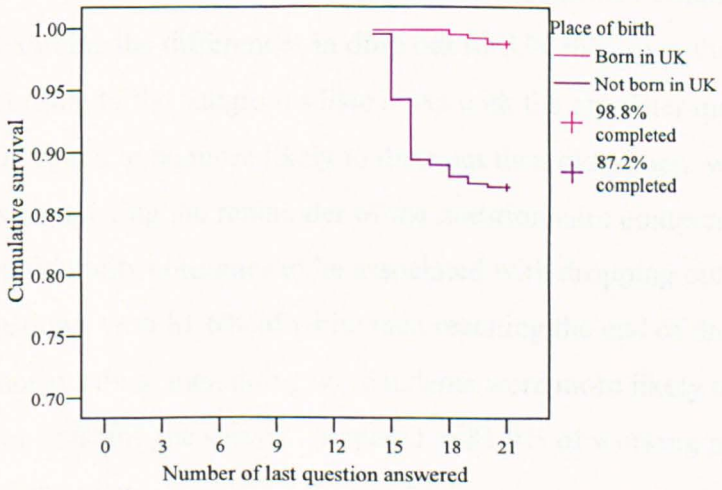


Figure 5.9: Pre-filter drop out 2003 by place of birth



The high drop out of men not living in the UK and born outside the UK occurred around questions 15, 16 and 17 which asked about country of birth, how long the man had lived in the UK and why he came to the UK. During the course of these questions, men who had already stated that they lived outside the UK were asked whether they were born in the UK and, if not, they were first asked their country of birth and then inadvertently mis-routed to one or two irrelevant questions about how long they had lived in the UK and why they came here. This combination of repeated questioning about place of birth and mis-routing appears to have prompted them to drop out.

5.3.3 Drop out over post-filter questions - UK men

The analysis up to this point in Chapter 5 has been concerned with all the men who originally clicked into the survey, including those living outside the UK. The focus of Part II of this thesis is on UK men and the analysis of post-filter drop out will therefore begin with a detailed description of the findings for this sample. This will be followed by a brief summary of the findings for London men, non-London men, UK men and all men who started the 2003 web survey.

It should be noted that the sample size for the comparison of drop out according to HIV status in the post-filter part of the survey is smaller than for other subgroup comparisons because questions about HIV status came just after the filter questions.

Table 5.2 (on the following page) shows the results of applying the log rank statistic to examine the differences in drop out for UK men over the post-filter questions according to the subgroups listed. As with the pre-filter questions, younger men were found to be more likely to drop out than older men, with only 76.3% of up to 24s completing the remainder of the questionnaire compared to 89.0% of the 45 to 54s. Ethnicity continues to be associated with dropping out over the post-filter questions, with 81.6% of white men reaching the end of the survey and 67.1% of minority ethnic men doing so. Students were more likely to drop out, with 75.1% of them finishing the survey compared to 81.9% of working men and 80.9% of non-working men.

Table 5.2: Log rank comparison of drop out during post-filter questions – UK men

	Log rank statistic
Website of origin	3.7
Week of completion	6.0
Time of day	0.4
Age group	26.8 **
Ethnicity	20.3 **
Work status	12.2 **
Education	9.3 **
Place of residence (London or outside London)	0.1
Place of birth	0.3
Type of area	4.2 *
Sexual orientation	27.1 **
Openness	26.3 **
Sex with only men	26.8 **
HIV status	7.3 *
Number of drop outs	*597*

Note: *$p \leq 0.05$, **$p \leq 0.01$

Not having a degree was also associated with dropping out (79.2% vs 83.5% finished). There were no significant differences according to whether respondents lived in London or outside the capital but men who lived in an urban area were more likely to drop out than rural-dwellers (80.2% vs 85.4% finished).

Dropping out over the main body of the survey was also associated with self-identifying as bisexual rather than gay (73.0% vs 82.6% finished), being less open about sexuality (72.2% vs 82.5% finished), having sex with both men and women (72.4% vs 82.5% finished) and having never tested for HIV, with HIV positive men the most likely to finish (84.7%), followed by HIV negative men (83.8%) and never tested (80.3%). These differences are illustrated in Figures 5.10 to 5.18 (on the following pages).

Figure 5.10: Post-filter drop out 2003 by age group - UK men

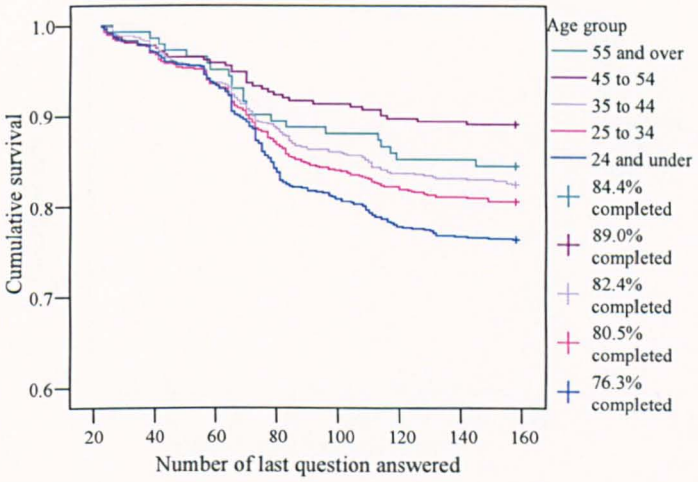


Figure 5.11: Post-filter drop out 2003 by ethnicity - UK men

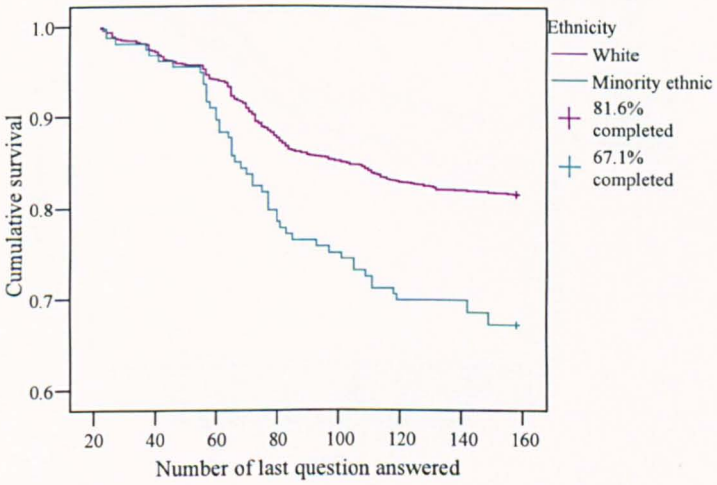


Figure 5.12: Post-filter drop out 2003 by work status - UK men

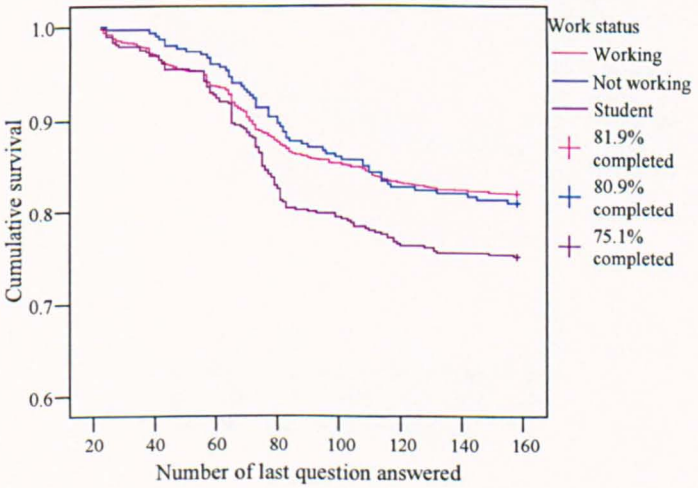


Figure 5.13: Post-filter drop out 2003 by education - UK men

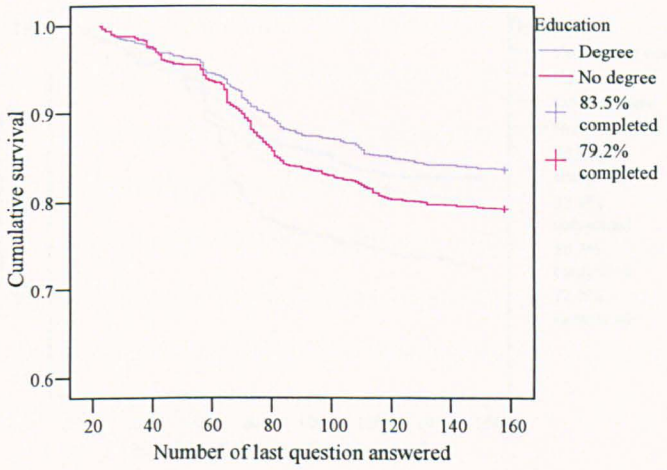


Figure 5.14: Post-filter drop out 2003 by type of area - UK men

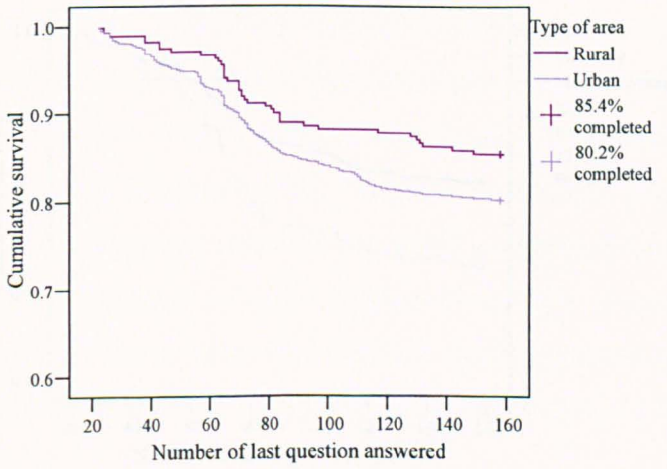


Figure 5.15: Post-filter drop out 2003 by sexual orientation - UK men

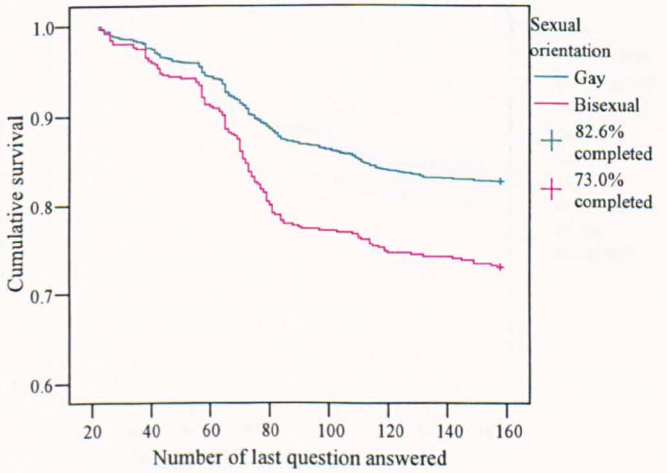


Figure 5.16: Post-filter drop out 2003 by openess - UK men

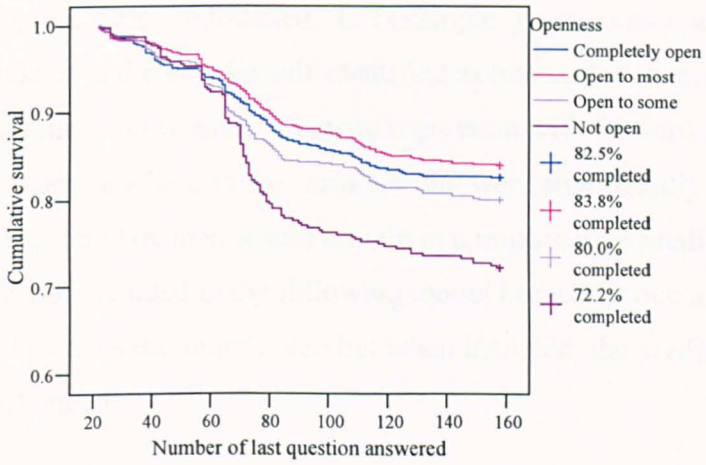


Figure 5.17: Post-filter drop out 2003 by sexual partners - UK men

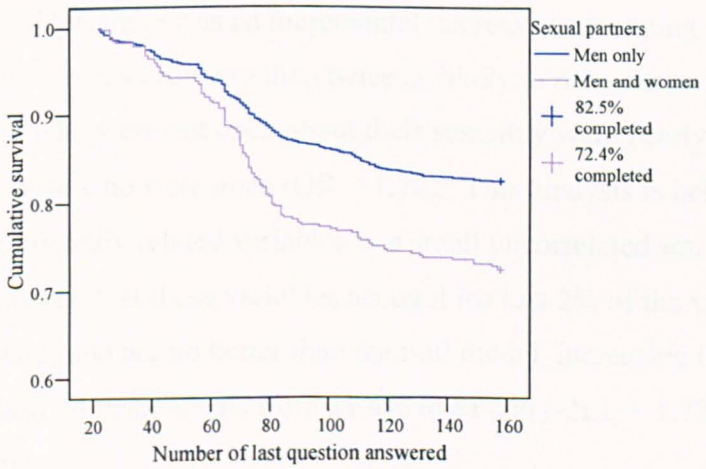
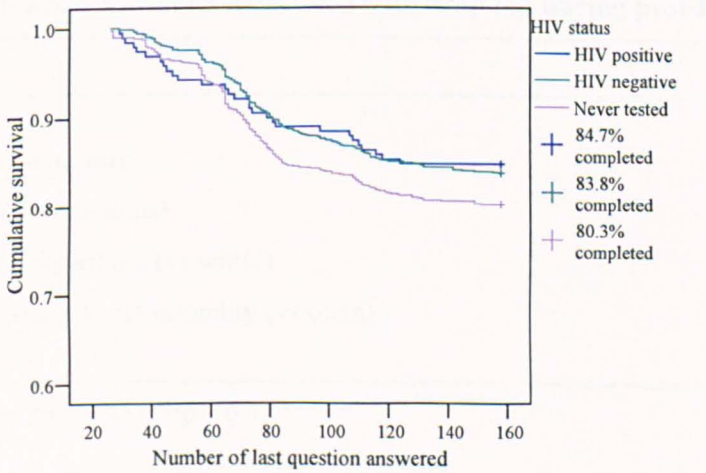


Figure 5.18: Post-filter drop out 2003 by HIV status - UK men



A number of the variables that were significantly associated with post-filter survey drop out were confounded. For example, younger men were more likely to be students and men who self-identified as bisexual were more likely to have had sex with men and women. Logistic regression with forward stepwise selection, was used to examine which of the variables that were significantly associated with dropping out among UK men would remain in a multivariate predictive model. HIV status was not included in the following model because it occurred after the filter questions and reduces the sample size but when included, the predictive variables remain unchanged.

Table 5.3 illustrates how being younger, minority ethnic and less open about orientation were significant predictors of drop out in the UK sample. As respondents got older, there was an incremental decrease in dropping out (OR = 0.98). Minority ethnic men were more than twice as likely to drop out as white men (OR = 2.25) and men who were not open about their sexuality were nearly twice as likely to drop out as those who were open (OR = 1.70). This analysis is helpful in reducing the list of significantly related variables to a small uncorrelated set. The results tells us, however, that these variables account for just 2% of the variation in dropping out (R^2 = 0.02) and are no better than the null model, increasing the percentage of correctly classified respondents from 81.4% to 81.5% (-2LL = 1,721.7, χ^2 = 39.6, 3df, $p \leq 0.01$).

Table 5.3: Variables associated with drop out during post-filter questions – UK men

	Odds ratio
UK men only	
Age (continuous)	0.98 **
Minority ethnic (vs white)	2.25 **
Not open about sexuality (vs open)	1.70 **

Note: *p ≤ 0.05 , **p ≤ 0.01

5.3.4 Drop out from 2002 web survey

An analysis of drop out of UK men from the Internet and HIV 2002 web survey was also undertaken, as detailed in Appendix 4. Even though all questions were compulsory in the 2002 web survey, the drop out patterns were very similar to those described above for the 2003 survey. Post-filter questions were completed by 79.7% of men in 2003 and 79.5% in 2002 but there was one notable subgroup difference. In 2003, HIV positive men were less likely to drop out over the post-filter questions than HIV negative men or those who had not tested for HIV whereas in 2002, HIV positive men were more likely to drop out than these other two groups.

Table 5.4: Log rank comparison of drop out during post-filter questions – all men

	London men	Non-London men	Non-UK men	All men
Website of origin	0.4	3.4	0.0	7.3 **
Week of completion	3.1	5.2	1.9	9.0
Time of day	0.4	0.3	4.6	0.6
Age group	13.0 *	19.0 **	14.1**	33.6 **
Ethnicity	22.3 **	2.9	2.2	28.1 **
Work status	5.8	10.1 **	2.1	15.1 **
Education	3.4	6.1 *	3.3	8.9 **
Place of residence	n/a	n/a	n/a	23.3 **
Place of birth	0.0	0.2	1.6	7.0 **
Type of area	n/a	4.9 *	n/a	4.2 **
Sexual orientation	0.9	27.6 **	3.1	31.0 **
Openness	7.1	21.1 **	0.0	24.6 **
Sex with only men	7.6 **	19.8 **	1.8	28.8 **
HIV status	3.3	5.6	4.6	8.0 *
Number of drop outs	*129*	*468*	*104*	*701*

Note: *$p \leq 0.05$, **$p \leq 0.01$

5.3.5 Drop out over post-filter questions – all men

Table 5.4 shows significant differences in subgroup drop out for the London, non-London, non-UK and all men. The log rank statistic was used to test the differences in drop out and is shown for each sub-sample comparison.

Among London men, those who were younger, from an ethnic minority and had sex with both men and women were significantly more likely to drop out over the post-filter questions. For non-London men, drop outs were significantly more likely to be found among those who were younger, students, not educated to degree level, lived in an urban area, identified as bisexual, were less open about their orientation and had sex with men and women. The only significant difference non-UK men was by age group, with increased age associated with greater likelihood of finishing the survey. The only non-significant differences for all of the men who clicked into the survey were according to week of completion and time of day, with all differences in the direction of those reported above for the other 2003 samples.

5.3.6 Influence of survey routing – UK men

Although there were 158 questions in the questionnaire, there was a substantial amount of question routing and respondents answered a greater or lesser proportion of the total, depending on how they responded to particular routing questions. It is therefore possible that the above differences between subgroups were prompted by the different number of questions with which various groups of respondents were presented, rather than more directly related to the characteristics of the subgroups themselves. The hypothesis would be that routing to a greater number of questions increases drop out. In order to examine this, the following will explore subgroup responses to a number of key routing questions.

Seven dichotomous variables were selected for examination based on the questions listed in Table 5.5 (on the following page). These questions came at the beginning of sections in the questionnaire where responding "yes" resulted in routing to further related questions and responding "no" resulted in skipping the set of related questions.

Table 5.5: Significant differences in subgroup response to routing questions

Question number	1	2	3	4	5	6	7
Starting from gaydar (vs gay.com)	-	yes**	no*	-	-	yes*	-
Earlier survey participation (vs later)	-	-	-	-	-	-	-
Time of day	-	-	-	-	-	-	-
Older (vs younger)	-	-	yes**	-	-	-	no*
White (vs minority ethnic)	-	-	-	-	-	-	-
In work or not working (vs student)	yes*	-	yes**	-	-	-	-
Holding a degree or equivalent (vs not)	no**	yes*	-	-	-	-	-
Living in a rural area (vs urban)	-	-	-	-	-	-	no**
Born in the UK (vs not)	-	-	-	-	-	no**	-
Identifying as gay (vs bisexual)	yes**	-	no*	-	-	yes**	yes*
Open about sexual orientation (vs not)	yes**	no*	no**	-	-	yes**	yes**
Sex with men only (vs men and women)	yes**	-	no*	-	-	yes**	yes**
Having tested for HIV (vs not)	yes**	-	yes*	-	-	yes**	yes*

Note: *$p \leq 0.05$, **$p \leq 0.01$

1. In the past 3 months have you had anal sex without a condom?
2. In last 3 months have you actually met a man face-to-face whom you found on the Internet?
3. In last 3 months have you had sex with a man you met through the Internet?
4. In last 12 months have you actually met a man face-to-face whom you found on the Internet?
5. In last 12 months have you had sex with a man you met through the Internet?
6. In last 3 months have you had sex with a man you met anywhere other than on the Internet?
7. In last 12 months have you had sex with a man you met anywhere other than on the Internet?

If being routed to a greater number of questions was the driving factor behind the above subgroup differences in dropping out, men who dropped out should be more likely to answer "yes" to these routing questions because this would result in the presentation of a greater number of questions. Similarly, men who finished the survey should be more likely to answer "no" to the questions. The following tests the null hypothesis that there would be no difference between subgroups of men on

their likelihood of answering "yes" or "no" to these seven questions. Because this analysis is at the group level, it is subject to ecological fallacy which occurs when a relationship between the characteristics of a group is used to infer that the same relationship applies to individuals within the group (Singleton & Straits 1999). Any inferences to individual behaviour are therefore necessarily speculative.

Table 5.5 shows where a significantly higher proportion of respondents in the subgroup listed were more likely to answer "yes" or "no" to each of the seven questions. The empty cells are where no significant differences were found in the likelihood of subgroups answering "yes" and "no". The subgroup categories identified are those that were more likely to finish the survey, with the exception of website of origin, week of completion, time of day and place of birth where no significant differences in drop out were found. They should therefore be more likely to answer "no" to these questions. Inspection of the table indicates that this was not so and the following will describe the findings in detail.

For the first question, "In the past 3 months have you had anal sex without a condom?", there were no significant associations between whether the respondent had had sex without a condom and website of origin, week of survey completion, time of day, age, ethnicity, type of area or place of birth. In support of a hypothesis that answering "no" to this question is associated with less likelihood of drop out, men who held degrees were less likely to drop out and were also less likely to have had recent sex without a condom (42.4% vs 47.5%, $p \leq 0.01$). Contrary to the hypothesis, men who were not students were less likely to drop out but were more likely to answer "yes" to this question (46.1% vs 40.4%, $p \leq 0.05$), as were gay men (48.3% vs 30.4%, $p \leq 0.01$), men who were open about their orientation (46.9% vs 32.4%, $p \leq 0.01$), men who had had sex with only men (47.9% vs 33.3%, $p \leq 0.01$) and men who had tested for HIV (52.3% vs 37.1%, $p \leq 0.01$).

Although there were no significant differences in drop out according to the respondent's website of origin, men who started from gaydar were more likely to answer "yes" to the question, "In last 3 months have you actually met a man face-to-

face whom you found on the Internet?" (73.1% vs 65.3%, $p \leq 0.01$). In support of the hypothesis, men who were open about their orientation were less likely to drop out and more likely to answer "no" (29.1% vs 23.3%, $p \leq 0.05$) and contrary to the hypothesis, men who held degrees were less likely to drop out but more likely to answer "yes" to this question (73.8% vs 70.1%, $p \leq 0.05$).

There were a number of significant differences for the third question, "In last 3 months have you had sex with a man you met through the Internet?". In support of the hypothesis, men who identified as gay were more likely to answer "no" (13.7% vs 8.6%, $p \leq 0.05$), as were men who were open (13.7% vs 7.2%, $p \leq 0.01$) and men who had sex with only men (13.2% vs 8.6%, $p \leq 0.05$). However, older men were less likely to drop out but the mean age of those who answered "yes" was significantly higher (33.5 vs 30.1 years, $p \leq 0.01$). Non-students who were also less likely to drop out and more likely to answer "yes" (88.2% vs 79.8%, $p \leq 0.01$), as were men who had tested for HIV (88.6% vs 85.0%, $p \leq 0.05$). There were no significant differences in drop out by website of origin but men from gaydar were more likely to answer "no" (13.6% vs 8.7%, $p \leq 0.05$),

There was no significant differences for the fourth and fifth questions about meeting or having sex with a man from the Internet over the past year.

The hypothesis was supported by two of the significant results for the sixth and seventh questions, "In last 3 months have you had sex with a man you met anywhere other than on the Internet?" and, "In last 12 months have you had sex with a man you met anywhere other than on the Internet?". Thus, older men were less likely to drop out and the mean age of those who had had sex with a venue partner in the past 12 months was also significantly lower (31.4 vs 33.1 years, $p \leq 0.05$) and men who lived in a rural area were less likely to drop out and less likely to have had sex with a venue partner in the past 12 months (14.5% vs 26.5%, $p \leq 0.01$). Men who came from gaydar were more likely to have had sex with a venue partner in the past 3 months although there were no significant differences in drop out by website of origin (59.0% vs 53.2%, $p \leq 0.05$) and men who were born in the UK were less likely

to have had sex with a venue partner in the past 3 months although there were no significant differences in drop out by place of birth (56.3% vs 66.5%, $p \leq 0.01$).

All other significant findings for these two questions were contrary to the hypothesis. Gay men were less likely to drop out but more likely to have had sex with men from other venues both in the past 3 months (60.7% vs 40.2%, $p \leq 0.01$) and in the past 12 months (27.7% vs 20.0%, $p \leq 0.05$); as were men who were open their orientation (3 months: 60.4% vs 34.1%, $p \leq 0.01$; 12 months: 28.0% vs 15.2%, $p \leq 0.01$), men who had sex with only men (3 months: 59.3% vs 47.3%, $p \leq 0.01$; 12 months: 28.1% vs 17.6%, $p \leq 0.01$) and men who had tested for HIV (3 months: 65.4% vs 47.8%, $p \leq 0.01$; 12 months: 28.9% vs 23.1%, $p \leq 0.05$).

The variables used in the above analysis are fairly crude indicators of the number of questions answered and, as previously described, these findings are at the group level and may not hold at the individual level. However, they suggest that being routed towards fewer questions was not associated with greater likelihood of completing the survey. In most cases subgroup members were similarly likely to answer "yes" or "no" to the routing questions, although there was some evidence that some of the respondents who were routed towards answering more questions may have been more likely to complete it. Thus, men who were gay, open about their sexuality, had sex with men only and had tested for HIV were less likely to drop out and were more likely to answer "yes" to three of the seven routing questions. It is unlikely that more questions *per se* led to a greater likelihood of completion but suggested that this is mediated by the somewhat different motivations of these subgroups and their engagement with the survey content.

5.3.7 Analysis of drop out pop-up

338 respondents (24% of the total number of drop outs) completed the drop out pop-up when they dropped out of the 2003 web survey. They could tick more than one reason for dropping out, but Table 5.6 (on the following page) shows that the majority of respondents gave just one reason for doing so.

Table 5.7 shows the reasons they gave, in order of frequency. The most popular reason by far was that the respondent did not have time to complete the survey, which 44.1% of those who completed the drop out pop-up included as a reason for dropping out.

Table 5.6: Number of reasons given for dropping out of 2003 web survey

	Frequency	Percentage
1 reason	287	84.9%
2 reasons	32	9.5%
3 or more reasons	12	5.6%
Total	*338*	*100.0%*

Table 5.7: Reasons given for dropping out of 2003 web survey

	Total number of ticks	Percentage of respondents
Don't have time	149	44.1%
Questions are boring	54	16.0%
Don't feel like it any more	40	11.8%
Don't think will be confidential	37	10.9%
Don't like last question	30	8.9%
Not relevant to self	29	8.6%
Questions too personal	24	7.1%
Not alone any more	24	7.1%
Having technical problems	19	5.6%
Topic not important	12	3.6%
Base	*338*	

Figure 5.19: Frequency of drop out because of time

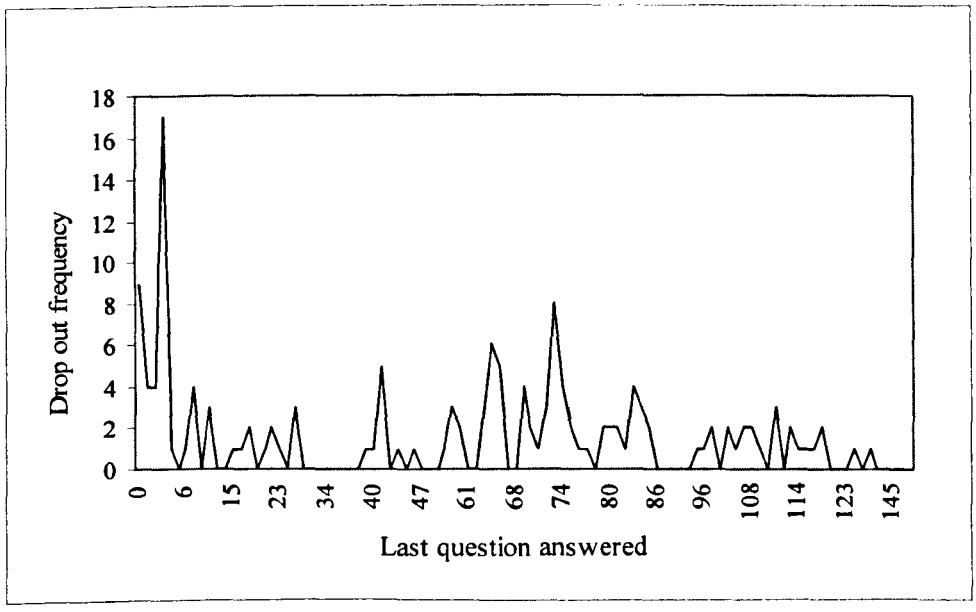
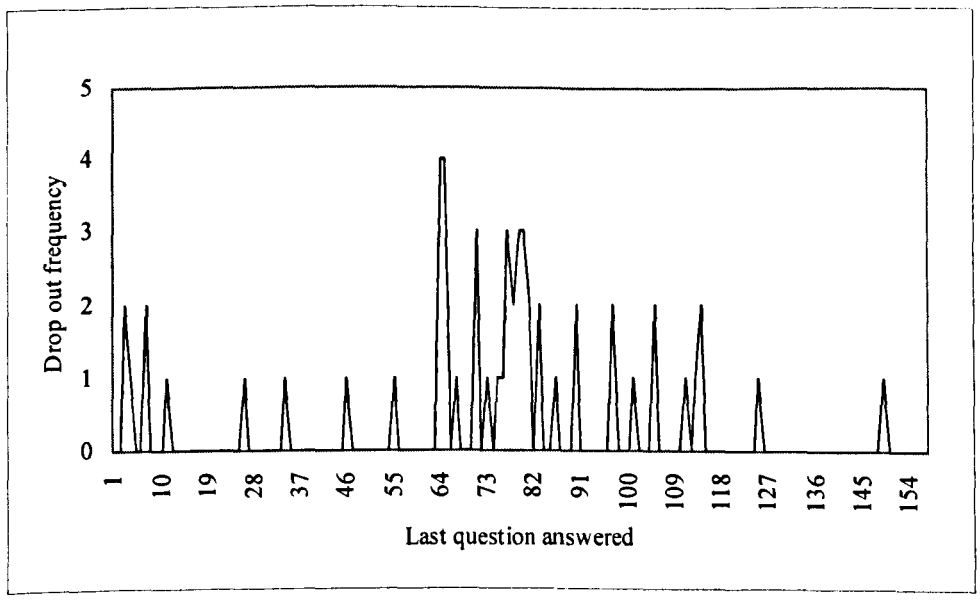


Figure 5.20: Frequency of drop out because of boring questions



Figures 5.19 and 5.20 identify the last question answered by respondents who said that they were dropping out because they did not have time to complete the survey or because the questions were boring. Figure 5.19 shows that the most time-related drop out occurs at question 3, which was about employment in the last week, where

17 respondents dropped out, and that following this, time-related drop out is fairly level, although it peaks again somewhat towards the middle of the survey. Figure 5.20 indicates that respondents were most likely to get bored with the questions in the middle of the survey and that having completed this section, boring questions were less likely to cause drop out.

Because confidentiality and question sensitivity are two of the key issues in surveys of sexual behaviour, drop out that occurred because the respondent did not believe that his answers would be confidential, because he did not like the last question or because he felt that the questions were too personal was also examined. Drop out because of confidentiality concerns occurred over the whole course of the survey, with the highest drop out of this type occurring before answering the first question, where four respondents gave this reason for quitting. Half of the drop out because of not liking the last question occurred in the sociodemographic questions at the beginning of the survey. About two thirds of this occurred during questions about work and job titles, with most of the remaining third occurring during questions about living in the UK. Drop out because of the personal nature of the questions was also spread over the survey and no particular questions appeared to generate this response more than others.

5.3.8 Analysis of drop out questions in motivational section

Questions were also included in the motivational section to collect data on reasons for dropping out. Of the 1,366 UK men who completed the motivational section, 26.4% (356) had seen the pop-up or banner before, 62.2% (840) had not seen the pop-up or banner before and 11.4% (154) were not sure. Of those men who had seen the pop-up or banner before, 64.2% (228) did not participate at that time, 6.8% (24) started the survey, 25.4% (90) finished it[3] and 3.7% (13) were not sure. They could give more than one reason for not starting the survey but again, the majority of respondents gave just one (Table 5.8 on the following page).

[3] It is possible that some of these men had completed a different online survey of sexual behaviour or the 2002 Internet and HIV web survey as described on p133.

Table 5.8: Number of reasons given for not starting 2003 web survey

	Frequency	Percentage
No reason	7	3.1%
1 reason	173	75.9%
2 reasons	47	20.6%
3 reasons	1	0.4%
Total	*228*	*100.0%*

Table 5.9: Reasons given for not starting the 2003 web survey

	Total number of ticks	Percentage of respondents
Didn't have time	147	64.5%
Didn't feel like it	101	44.3%
Someone joined me	6	2.6%
Didn't think would be confidential	4	1.8%
Thought would be boring	3	1.3%
Topic not important	3	1.3%
Thought not relevant to self	2	0.9%
Questions too personal	2	0.9%
Was having technical problems	2	0.9%
Base	*228*	

Table 5.9 gives the reasons that men gave for not starting the survey before, in order of frequency. The data show that not having time to complete the survey was the most popular reason. As these were highly motivated men who eventually completed the full questionnaire and the motivational section, it is not surprising that issues of confidentiality, boring questions, survey topic, survey relevance and

personal questions were not important to them. It is perhaps more interesting to note that situational factors were also unimportant, with only 2 men reporting technical problems and 6 put off through interruption by another person.

Only 24 men said that they had started the survey on a previous occasion but not finished it (Table 5.10). Again time was the most important reason for dropping out but technical problems were a somewhat more important issue for men who had dropped out and then come back to complete the survey at a later date.

Table 5.10: Retrospective reasons given for dropping out of the 2003 web survey

	Total number of ticks	Percentage of respondents
Didn't have time	14	58.3%
Had technical problems	5	20.8%
Didn't feel like it	2	8.3%
Didn't like a particular question	2	8.3%
Someone joined me	1	4.2%
Topic not important	1	4.2%
Thought not relevant to self	1	4.2%
Didn't think would be confidential	0	0.0%
Thought was be boring	0	0.0%
Questions too personal	0	0.0%
Base	*24*	

5.4 Discussion of findings

This discussion begins with a summary of the overall drop out from the Internet and HIV 2003 web survey and goes on to explore drop out in the early stages and late stages of the survey. It provides a brief summary of the reasons respondents gave for dropping out before concluding with the implications for researchers.

In 2003, respondents clicked into the survey 4,271 times and completed it on 64% of occasions. Although a wide range of completion rates has been reported in the literature (Musch & Reips 2000), this completion rate is consistent with a comparable web survey of sexual behaviour among a similar population of men in the USA (Ross *et al.* 2004).

5.4.1 Drop out in the early stages

A high proportion of the drop out occurred in the early stages of the web survey which is also consistent with previous findings (Jeavons 1998; Knapp & Heidingsfelder 2001; Lindley *et al.* 2003; Ross *et al.* 2003; 2004). The median last question answered was 20. Much of the drop out therefore occurred during the sociodemographic questions at the beginning of the questionnaire. Although such questions have been implicated in drop out (Jeavons 1998), this finding suggests that the position of a question within the web survey may be more important than question content in prompting drop out. The potentially sensitive nature of the main part of the survey did not result in increased drop out, in the same way that dropping out of a telephone survey of sexual behaviour was unrelated to question content (Catania *et al.* 1996). Rather, it seems that the principle of consistency, described in Chapter 4, comes into play. Once respondents have completed the early section of the survey, this norm dictates that they continue with what they have started and become less likely to drop out.

Closer inspection of the data reveals a high level of drop out by the non-UK men over questions 15, 16 and 17. Although these men were not included in the analysis for the Internet and HIV study, the inadvertent mis-routing of non-UK men to irrelevant questions appears to have prompted them to drop out. As the Internet and

HIV study was only concerned with London men, the web survey was not piloted on non-UK men, which may explain why the mis-routing went unnoticed. The finding illustrates how mis-directing respondents to questions that are not applicable to them may cause drop out, even when such questions may be skipped. Careful piloting of web surveys is needed to ensure that this does not occur.

Respondents who took part in the first week that the web survey was fielded were more likely to complete the pre-filter questions and continue to the main part of the survey. This concurs with Jeavons' (1998) finding that surveys that are fielded for a shorter time result in lower drop out. More regular use of the website to which the survey is attached has also been associated with a higher likelihood of completion (Knapp & Heidingsfelder 2001) and is likely to be behind the present finding. Although drop out according to the frequency of using sites like gaydar and gay.com could not be directly examined (because this question came too late in the questionnaire), there is a significant association between logging on to these sites and responding in the first week which corroborates this proposition.

Younger men and minority ethnic men were more likely to quit the web survey in the early stages. Although this raises the problem of sample bias, the reduced size of the minority ethnic sample is an additional source of concern for this group. Only 5.1% of the UK men who answered the ethnicity question in the 2003 web survey were from an ethnic minority and, having started the survey, these men were more likely to drop out which may further weaken the statistical power to detect differences according to ethnic origin.

5.4.2 Drop out in the later stages

The drop out pattern for subgroups of UK men over the post-filter questions bore certain similarities to those found in similar web surveys of sexual behaviour (Ross *et al.* 2003; 2004). As with the sample of UK men analysed here, Ross *et al.* (2003) found that men who were more highly educated were more likely to complete their web survey. Their sample of gay and bisexual men were also more likely to complete the survey than their sample of heterosexual men. In a subsequent study of

Latino MSM, Ross *et al.* (2004) found that gay men were less likely to drop out than bisexual or heterosexual men. A similar analysis of the Internet and HIV web survey also shows that gay men were more likely to complete than bisexual men. Being more open about orientation and having sex exclusively with men were also associated with completion.

Urban-dwelling UK men were more likely to drop out over the post-filter questions of the Internet and HIV 2003 web survey whereas they were less likely to drop out of Ross *et al*'s (2003) general population survey. Younger UK men were also more likely to drop out but less likely to do so in Ross *et al*'s (2003) general population survey and no age differences were found among the Latino MSM (Ross *et al.* 2004). These inconsistencies require further exploration, particularly in relation to the impact of age on response to questions of a sexual nature among gay and bisexual men.

Minority ethnic UK men continued to be more likely to drop out over the post-filter questions. Minority ethnic respondents were also more likely to provide incomplete answers to the self-completion questions of Natsal 1990 (Kupek 1998). This suggests that drop out and item nonresponse may be driven by similar factors and the relationship between ethnicity and item nonresponse will be explored in Chapter 7.

The drop out over the post-filter sections of the 2003 survey reflects the findings from the Chapter 4 on motivation for participation. The same subgroups who were more likely to be motivated by an interest in the survey topic were, on the whole, those subgroups that were more likely to complete the survey. Thus, men who were older, non-students, self-identified as gay, more open about their sexuality and HIV tested were more likely to be motivated by the survey topic and less likely to drop out. Engagement in the survey topic, which has been shown to be closely linked to the altruistic desire to contribute to academic research, appears to be a key factor in survey completion as suggested by Ross *et al.* (2004). The analysis of the influence of survey routing also indicates that answering more questions was associated with a

lower likelihood of drop out, suggesting that where the survey content is of greater relevance to the self, the respondent is less likely to lose interest.

Confining the analysis to London men had a notable impact, with significant differences remaining for just three variables - age, ethnicity and sex with men. This tends to indicate that there was a greater conformity among the response behaviour of London men.

Despite these differences, the independent variables used here accounted for very little of the overall variation in whether or not UK men were likely to drop out of the 2003 web survey. Although these variables have been shown to have an impact on drop out, a great deal of what happens when a respondent completes a survey of sexual behaviour was not captured by this analysis and, as suggested by the literature review, drop out is likely to be influenced by a number of survey design factors that were not examined here.

The findings for drop out of UK men over the post-filter questions were very similar for the 2002 web survey. The only group displaying different response behaviour was HIV positive men. In 2003, they were the more likely to finish the survey than HIV negative or never tested men whereas in 2002, they were the less likely to do so. We can only speculate about what may have caused this difference. One of the main differences between the surveys was that every question in the 2002 survey was compulsory whereas only routing questions were compulsory in the 2003 survey. Anecdotal evidence from an HIV positive interviewee in the qualitative interviews described in Chapter 8 indicated that he was frustrated by a question where he felt that his response gave the false impression that he had sex with a man of unknown HIV status. Although he answered the question and went on to complete the survey, similar frustration may have driven other HIV positive men to drop out of the 2002 web survey where every question required an answer.

5.4.3 Drop out pop-up and questions

The reason most often given for dropping out was lack of time. This is a situational factor over which the researcher has no immediate control. Conrad *et al.* (2003) have shown that higher drop out occurs when the perceived progress is manipulated so that respondents feel they have made less progress through the survey. The higher degree of time-related drop out that occurs in the middle of the survey may similarly be due to respondents feeling that they still have a long way to go. It was also at this point that respondents said they lost interest because of the boring questions.

5.4.4 Implications for researchers

Drop out was most likely to occur at the beginning of the survey. Men who were younger, minority ethnic and not open about their sexual orientation tended to be more likely to drop out in the later stages but neither respondent characteristics nor specific questions of a sensitive nature were a major cause of post-filter drop out. This suggests that drop out is unlikely to introduce major systematic bias into the estimates derived from web surveys such as this. Respondents indicated that time was an important cause of drop out and, although the researcher has no control over this, every effort should be made to ensure that respondents are encouraged by their rate of progress. The negative effects of mis-routing suggest that the importance of thorough piloting should not be underestimated.

6 Comparison of Internet and offline samples

Synopsis

Secondary analysis was used to compare the characteristics of the men who participated in the Internet and HIV 2002 and 2003 web surveys with the MSM drawn from Natsal 2000. While the Internet national samples were similar to the Natsal national sample on a range of variables, there were marked differences between samples on variables indicating HIV risk behaviour, and greater differences between samples of London men.

6.1 Introduction

Chapters 4 and 5 have examined some of the factors that are likely to shape the final composition of Internet samples of gay and bisexual men. The present chapter focuses on the actual composition of such samples. Its purpose is to compare the characteristics of the samples of gay and bisexual men who participated in the Internet and HIV web surveys with a sub-sample of the MSM who participated in Natsal 2000, a random probability sample of British adults. It seeks to answer the following question:

"How do self-selected Internet samples of gay and bisexual men differ from men who have sex with men drawn from a probability sample?"

The rationale concurs with what the authors of a paper on the first National Survey of Sexual Attitudes and Lifestyles, Natsal 1990, wrote:

> *"random sample survey data could provide a context for studies based on volunteer samples of individuals at high risk of HIV infection such as homosexual men and injecting drug users"* (Wadsworth *et al.* 1993: 407).

Researchers from Natsal 2000 also note that increases in reporting risk behaviour among gay and bisexual men in the 1990s were observed in both Natsal and gay

community surveys (Mercer *et al.* 2004), which is further reason for exploring why such similarities may occur. It is furthermore suggested that data from Internet samples should be validated against that from non-Internet samples (Hewson *et al.* 2003).

The introductory section to this chapter begins with a description of what is meant by probability and non-probability sampling and describes how these designs were applied to the Natsal probability sample and the online and offline convenience samples of the Internet and HIV study. It goes on to examine the probability and non-probability techniques that have been applied to sampling respondents for other Internet surveys and to review the current literature on the characteristics of respondents to Internet surveys. Before looking specifically at gay and bisexual men, it continues with a review of the literature on the characteristics of a more general population of respondents to surveys of sexual behaviour. The chapter then examines how researchers have sought to sample gay and bisexual men and the role that the Internet has played in recruiting samples from minority populations. It concludes with a summary of studies that have examined the characteristics of gay and bisexual respondents to Internet surveys.

6.1.1 Purpose of sampling

Rather than taking censuses of the whole population of interest, samples are drawn because they can be studied more quickly and cheaply. Available resources can then be concentrated on gathering high quality data from the sample and directed to other areas of the research process (Fink 2003; Lynn 1996a). At the same time, it is important that members of the sample should represent the population from which they are drawn in order to make well-founded inferences to the population of interest. In this way,

> *"the task of the research designer is to effect an appropriate balance between cost and accuracy"* (Lynn 1996a: 131).

Such accuracy is best obtained with a probability design, where each member of the population of interest has a known non-zero possibility of being included in the sample. Probability samples are expensive and time-consuming to implement, with the result that non-probability designs are frequently used in their place. Natsal 2000, for example, was budgeted to cost an estimated £1.85 million (C Mercer, personal communication, 5th December 2005) and fieldwork alone took nearly two years.

In any discussion of sampling, the issue of sample size is always paramount. A large sample has a number of advantages. For probability samples, the precision or standard error of the estimate is calculated as the standard deviation divided by the square root of the sample size. As sample size increases, therefore, the variance around the estimate decreases and its precision increases. The formula used to calculate sample size means that the relationship between sample size and precision is not linear and after a certain point of around 1,000 responses, the increase in precision becomes less pronounced and further increase in sample size may become uneconomical (Bryman 2004). Whereas increasing the size of a probability sample diminishes the variance, however, it may do nothing to diminish the bias of a non-probability sample (Kalton 1983). Such bias occurs when the sample differs systematically from the population which it is intended to represent and is a problem in many web surveys that are recommended for their large sample sizes but subject to the bias associated with surveying a volunteer sample of Internet users and generalising to a more broadly defined population.

Bryman (2004) identifies three further important elements for determining sample size. Probability samples should include an estimation of the expected level of non-response and any sampling design must consider the degree of heterogeneity in the population and the type of analysis to be used. Thus, where the researcher wants to examine the response of sample subgroups, the number of members in each subgroup should be sufficient for this analysis.

What follows is a brief discussion of how probability and non-probability designs may be implemented (for more detailed discussion see Kalton 1983) and how these principles were operationalised in the Natsal and Internet and HIV studies.

6.1.2 Probability sampling

The key feature of probability sampling is that each element of a sample selected from the survey population must have a known non-zero chance of being included. One of the first steps in probability sampling, then, is to define the population or *"totality of elements under study"* (Kalton 1983: 6). Although never simple, the potential complexity of this task was illustrated earlier in the discussion of the issues around defining the population of gay and bisexual men. Kalton (1983) goes on to define the target population as that which would be ideal to meet the survey objectives and the survey population as that which is modified in order to take into account the practical constraints of the survey, with exclusions from the target population explicitly identified.

In order to select a probability sample, it is essential that a complete listing of elements or sampling frame is available (Kalton 1983). Although social research is mainly concerned with sampling individual people, researchers may be interested in other units of analysis, such as households or particular establishments, for which a sampling frame is similarly required. The frame should contain no missing elements, no ineligible elements and no duplicates (Lynn 1996b). The main sampling frame used for national probability samples of individuals in the UK is the postcode address file (PAF), which is a computerised list of every address to which mail is delivered in the UK and is the recommended alternative. The small user PAF is used to sample residential addresses and, although it contains some non-residential property, just one or two percent of residential addresses are missing (Lynn 1996b). Another sampling frame that is popular in the United States consists of all known area code telephone numbers from which complete telephone numbers are randomly generated. Sampling from this frame is known as random digit dial or RDD.

The use of PAF and RDD illustrate one of the problems of sampling which is coverage. People who live in places such as prisons, care homes or student halls of residence are excluded from the sampling frame when PAF is used and people who do not have a telephone are excluded when RDD is used. In this way, elements from the target population may be missing and unless an appropriate survey population can be defined taking this into account, the sample may be subject to non-coverage bias.

Table 6.1: A taxonomy of probability selection methods

I. *Epsem*: equal probability for all elements (a) Equal probabilities at all stages (b) Equal overall probabilities for all elements obtained through compensating unequal probabilities at several stages	*Unequal probabilities* for different elements; ordinarily compensated with inverse weights (a) Caused by irregularities in selection frames and procedures (b) Disproportionate allocation designed for optimum allocation
II. *Element Sampling:* single stage, sampling unit contains only one element	*Cluster Sampling:* sampling units are clusters of elements (a) One-stage cluster sampling (b) Subsampling or multistage sampling (c) Equal clusters (d) Unequal clusters
III. *Unstratified Selection:* sampling units selected from entire population	*Stratified Sampling:* separated selections from partitions, or strata, of population
IV. *Random Selection* of individual sampling units from entire stratum or population	*Systematic Selection* of sampling units with selection interval applied to list
V. *One-Phase Sampling:* final sample selected directly from entire population	*Two-Phase (or Double) Sampling:* final sample selected from first-phase sample, which obtains information for stratification or estimation

Source: Survey Sampling (Kish 1967: 20)

Kish (1967) devised the taxonomy of probability selection methods that is presented in Table 6.1. It describes five alternatives that may be combined in probability sample designs.

Natsal 2000 defined its target population as men and women aged 16 to 44 living in private households in Great Britain (see Erens *et al.* 2001). The survey adopted a multistage stratified probability design. Postcode sectors from the small user PAF were selected in the first instance, representing clusters of addresses. Addresses were then selected from these sectors, representing households which may form clusters of individuals. At addresses containing more than one eligible adult, one person was randomly selected using a Kish grid technique. The sample was also stratified, with addresses in London over-sampled. This is an example of disproportionate allocation, because the number of addresses sampled from the London and non-London strata was not proportional to the number of addresses within and outside London. This design was adopted because the prevalence of many HIV risk behaviours is higher in London and the researchers wanted to be able to estimate this more precisely and to examine subgroups of those at risk from HIV transmission. A final sample of 12,000 respondents was attained from the original sample of nearly 40,000 addresses.

Natsal included an independently selected boost sample of black and Asian adults to address the issue of inequalities of health and the small proportion of ethnic minority respondents found in national general population surveys. This was also a multistage design, but involved "screening" households to establish whether they contained at least one eligible adult from a black or Asian ethnic group. Screening is one of the techniques that is used in sampling minority populations and will be described in more detail below. An ethnic boost of 735 black and Asian adults was thus attained from the sample of 15,000 addresses.

The above has described how samples are designed so that each element is randomly selected from the population of interest and, in order to maintain the representativeness of such carefully selected samples, it is important that every effort

is made to secure the participation of sample members. A response that is not obtained is known as unit non-response and results in non-response bias to the extent that non-respondents differ systematically from respondents.

6.1.3 Non-probability sampling

Non-probability samples can be drawn in various ways. The methods known as convenience sampling, purposive sampling, quota sampling and snowball sampling will be examined here. The idea with ***convenience sampling*** is that the researcher recruits respondents from a group of individuals that is readily available (Bryman 2004; Fink 2003), creating what may be seen as an opportunistic and voluntary sample. Because of the lack of control that the researcher has over who participates in a convenience sample, the use of such samples has been described as making,

> *"educated guesses about the nature of the population"* (Best *et al.* 2001: 133).

Such "educated guesses" may be the best way forward in researching particular populations, however, supporting an argument that it is better to undertake research with recognised limitations than to ignore such topics because of the methodological problems (Koch & Emrey 2001). This has proved to be the case in research involving the sexual behaviour of gay and bisexual men which has generally relied on convenience samples (Kalton 1993; Martin & Dean 1993; Rhodes *et al.* 2002). Indeed it is argued that,

> *"a random selected study of men who have sex with men (MSM) is difficult, if not impossible, to achieve"* (Tikkanen & Ross 2000: 614).

The issues behind this will be discussed in *6.1.6 Sampling minority populations.*

In common with other surveys of gay and bisexual men, the Internet and HIV study adopted a convenience approach to sampling (see Elford *et al.* 2004a). One of the strengths in the design of this study, however, was that it sampled gay and bisexual London men from four different sources. The Internet and HIV study is described in

detail in Chapter 2 but, to summarise, the Internet samples that will be compared to the Natsal sample of MSM were recruited through pop-ups and banners appearing in the two most popular websites for gay men in the UK and the remaining offline samples were recruited in London gyms, in an outpatients clinic for HIV positive men and in HIV testing clinics.

A second non-probablity method is known as ***purposive sampling***. In this case, the researcher attempts to control for variation in the population of interest by using his or her expertise to select a sample which is deemed to represent the population. Convenience and purposive samples are often combined into ***quota sampling*** (Lynn 1996a). Like probability sampling, quota sampling actively seeks to produce a sample that represents the population of interest and purposively selects it on this basis. The sample is selected on a quota system whereby specified numbers of respondents are recruited to fit certain criteria. The selection of these quotas of respondents is not done on a random basis, however, but it is left to the interviewer to approach individuals who happen to be available at the time and place where they are undertaking data collection in order to fulfil their quotas. The subjective nature of this selection process means that researcher cannot guard against the possibility of sample bias.

The final non-probability method is ***snowball sampling***. Use of this method requires an initial identification of eligible persons who are then asked to identify other members of the population. The process of referral generally continues until the sample size is met. One of the problems associated with this type of sampling is that respondents are likely to be well connected to whichever population is being sampled and socially isolated members will have little or no chance of selection.

6.1.4 Online sampling

Although online sampling may be characterised along the traditional dichotomy of probability and non-probability methods, it has specific features that need to be considered. Couper (2000) defines a number of approaches to sampling for web

surveys, including both probability and non-probability methods. This model will be used here as the framework for discussing online sampling.

Couper (2000) lists five probability-based methods. The first is *intercept surveys,* whereby every *n*th visitor to a site is invited to participate via banner advertisements or pop-ups and the population can be defined as the number of visits or visitors to the site. The second is *list-based samples of high-coverage populations*, with the list typically made up of email addresses from which the sample is drawn. This approach is suitable for intra-organisational surveys or student populations. The third approach is *mixed-mode designs with choice of completion mode* where the respondent may choose to complete the survey by web or other mode. The fourth is *pre-recruited panels of Internet users* where respondents are recruited to the panel via probability sampling. The final method is a *probability sample of the full population* where respondents are recruited via probability sampling and provided with the necessary equipment if they do not already have it.

Couper lists three non-probability approaches. *Web surveys as entertainment* are included because the general public often confuses such surveys with scientific surveys. *Self-selected web surveys* are those whereby an open invitation is posted on a website or elsewhere and no control exists over who completes it or how many times. *Volunteer panels of Internet users* consist of panels of self-selected volunteers who are then sampled for inclusion in particular surveys.

6.1.4.1 Online probability sampling

The concept of the ***intercept survey*** is intuitively appealing because it appears to give an insight into the response rate for the survey through calculation of the proportion of responses as a percentage of the invitations to participate. Because the target population is narrowly defined as users of a particular website, however, its application is limited to consumer surveys of customer satisfaction or website evaluation (Couper 2000).

The use of ***list-based samples of high-coverage populations*** cannot be extended to the general population of Internet users because there is no comprehensive list of email addresses that would correspond to the post-code address file and there is no way of randomly generating email addresses that would correspond to RDD. Where lists of email addresses have been generated by other means, sampling from them can be problematic because email addresses are easily changed at any time (Smith 1997). At least 5% of emails were returned as undeliverable in one Internet survey and this was considered to be an underestimation of non-delivery rate because 23% of emails that were previously distributed to imaginary addresses failed to be returned as non-deliverable (Swoboda *et al.* 1997). Another issue with regard to the inaccuracy of email sampling frames is that individuals often have more than one email address. Furthermore, they may be reluctant to submit their email addresses to contact lists. Only 39% of paediatricians listed their email address in a professional directory although more than 90% of those surveyed had access to email (McMahon *et al.* 2003).

In spite of these issues, Internet researchers have attempted to generate lists of email addresses to use as sampling frames for a general population of Internet users. Sheehan and Hoy (1999) sent 5,000 solicitation emails to addresses drawn from an online directory with the opportunity to opt out. 1,276 (26%) of these addresses were undeliverable and 433 declined to participate and were replaced by reserve names. Of the 3,724 surveys distributed, 889 useable responses were returned, resulting in a sample that is unlikely to represent the members of the online directory, let alone the general population of Internet users.

Another attempt to generate a representative sample of Internet users involved random selection of two hundred of the eight thousand newsgroups that were available to the researchers (Swoboda *et al.* 1997). These were scanned for email addresses over a period of fourteen days, during which time nearly nine thousand addresses were identified. This yielded a worldwide response of 1,713 or 20% to their opinion poll of global issues, raising the possibility of non-response bias when 80% of the sample failed to respond and securing a sample that tended to be young

(mean age 30 years) and male (87%) with strong representation from the student population (25%). Although the study did not claim to generate a sample that was representative of the general population, we should still question the extent to which regular newsgroup users represent the population of Internet users and the non-coverage bias that may result. Certain special interest groups may be reached by targeting particular newsgroups, sites or discussion lists, but frames of email addresses generated from these sources can only represent a narrowly defined survey population.

List-based sampling is particularly effective, however, in surveys of organisations where pre-existing lists of email addresses for the survey population have high coverage and accuracy (Couper 2000). Where staff or students use email addresses provided by the organisation, for example, a comprehensive sampling frame may be drawn up and elements randomly selected from the list. A good example of how effective this method may be can be seen in the work of survey methodologists at Michigan University (McCabe *et al.* 2002). 7,000 undergraduate students were randomly assigned to either web-based or mail-based survey mode and participation was requested by email or postal invitation respectively. Not only was the response rate for the web survey significantly higher than for the postal survey (63% vs 40%), but gender composition of the web survey sample matched the target population more closely. A similar study which randomly assigned 600 students to a web or postal survey achieved similar response rates across modes (58% vs 62% respectively) and found no significant differences between mode in the respondent profile (Pealer *et al.* 2001).

Even where email lists have high coverage and accuracy, one of the main problems associated with use of email to recruit respondents is that unsolicited emails may be perceived as "spam" (Witmer *et al.* 1998) and have been found to annoy some respondents (Mehta & Sividas 1995; Smith 1997). Although only a minority of respondents may complain about unsolicited survey requests, the capacity to delete them is just a click away and they are likely to be automatically filtered out of the inboxes of increasingly sophisticated email users (Birnholtz *et al.* 2004). The

positive effect of personalising the survey request which has been long recognised in postal surveys (Dillman 1978; 2000), is also likely to apply to email survey requests. Thus, individually addressed emails were found to increase response rates over bulk email (Alves & Szucs 2001). Banner advertisements are also likely to violate privacy by intruding on the personal space with unwanted sights or sounds but to a much lesser extent than unsolicited emails (Cho & LaRose 1999). With regard to surveys of sexual behaviour, it may also be the case that sending out requests for participation is unethical because people often read email at work or share accounts with family members (Binik *et al.* 1999).

Probability samples that are not recruited via email might be approached using a different mode and given the choice of how they complete the survey. An interesting example of the ***mixed-mode design with choice of completion mode*** took place when the citizens of Thomasville, Georgia, were given the opportunity to participate in a survey relating to perception of community policing (Ballard & Prine 2002). The survey was mailed to every home, a total of nine and a half thousand, and advertised on the city's website. It achieved a response rate of 26%, with 71% of responses returned by post and 29% electronically. Although there was no modal difference in the response of males and females, the Internet sample was more likely to be white, younger, wealthier and college-educated. The two samples were more alike than different, however, with similar perceptions of community policing, leading the authors to conclude that both samples may be biased towards those who were likely to return such a survey albeit by post or by Internet.

It is possible to generate ***pre-recruited panels of Internet users*** by using probability methods to identify Internet users but such panels tend to be recruited through the non-probability methods discussed below. Although these panels limit generalisations to the population of Internet users, it is possible to generate a ***probability sample of the full population*** at the expense of providing the necessary computer equipment to all members of the sample. This type of sampling is only economical where panels are repeatedly surveyed and has been successfully adopted by the market research organisation Knowledge Networks which recruits its online

panel via random digit dialling and provides members with the Internet access needed to respond to their surveys. The panel may thus include respondents who were not Internet users at the time of recruitment.

This is important given that the population of Internet users is different from the wider population. For example, 15% of UK families in the lowest income group had home access to the Internet in 2003-04, compared to 89% of the highest income group (Office for National Statistics Expenditure and Food Survey 2003-04, http://www.statistics.gov.uk). Population estimates based on Internet samples are therefore likely to be biased to the extent that low income groups differ from high income groups, echoing the Literary Digest's mistaken choice of the winner of the 1936 presidential contest (Webb & Wybrow 1974).

6.1.4.2 Online non-probability sampling

Given the expense of probability sampling in general and the particular problems of online probability sampling outlined above, it is not surprising that online convenience samples are more commonly used than probability samples (Bradley 1999). Although Couper (2000) includes ***web surveys as entertainment*** in his model as a non-probability method, the literature examined for the purposes of this exploration of online sampling did not contain any references to them. A format that is particularly prevalent, by contrast, and often cited in the literature, is the ***self-selected web survey***. This is also the design used by the Internet and HIV web surveys and is therefore of particular interest here.

The first publicly accessible web-based survey was posted on the Internet for one month in January 1994 with the aim of exploring the profile of web users, their reasons for using the web and their opinions about it. (Kehoe & Pitkow 1996). The purpose and URL of this first GVU (Graphics, Visualization and Usability Center) WWW User Survey were announced in a computing newsgroup and on two high profile Internet-related websites (the National Center for Supercomputing Applications and the Centre Européen de Recherche Nucléaire) resulting in a response of 4,777 over the one-month period (Pitkow & Recker 1994). GVU

surveys were then conducted every six months and by the fifth survey, researchers were promoting it through diverse media, including various popular websites, newsgroups, the local and national press and a respondent mailing list (Kehoe & Pitkow 1996).

Another oft-cited example of a self-selected web survey is Survey 2000 which, with nearly 50,000 worldwide participants, claimed to be the largest and most comprehensive Internet-based social science survey of its time (Witte *et al.* 2000). Respondents accessed the survey from the National Geographic Society (NGS) homepage or picked up the URL from advertising in NGS periodicals. As with all self-selected survey samples, the issue of the representativeness of the sample needs to be addressed. The responses to Survey 2000 were weighted according to US national population statistics derived from the General Social Survey, the Panel Study of Income Dynamics and the National Longitudinal Survey of Youth. This type of weighting is known as post-stratification and aims to match the profile of the sample to the population profile where the proportions of characteristics of importance to the study do not correspond (Lynn 1996a). In this case, weighting by central demographic variables alone was not found to yield plausible generalisations to the population. This is an issue of key importance that will also be discussed in relation volunteer panels of Internet users and how researchers have attempted to deal with the problem of who responds to Internet surveys.

As described earlier, the sample composition is dependent on where the respondents were recruited and this is equally true of self-selected web samples (Hewson *et al.* 2003). Users of diverse Internet websites, newsgroups or discussion lists do not form a homogenous group and this is reflected in the samples that are attained when respondents are recruited from different places. For example, recruitment via banner advertisements resulted in a different profile of respondents to those asked to opt in as they completed online subscription forms (Alvarez *et al.* 2003). Schillewaert *et al.* (1998) also found that respondent profiles were dependent upon the method by which they were recruited. Newsgroup respondents were likely to be younger, on lower incomes and more longstanding, heavier Internet users than respondents

recruited via traditional media, other websites or an online database. It is therefore recommended that respondents are recruited from several Internet locations in order to diminish the bias associated with recruiting from one platform (Mustanski 2001) although this cannot eliminate the potential bias associated with convenience sampling. The Internet and HIV web survey recruited respondents via two websites, gaydar and gay.com, and it is therefore of interest to compare the composition of these two sub-samples.

One of the issues associated with open-access web surveys which is often raised, is that of multiple submissions. Concern about this resulted in the introduction of access codes at the second round of an organisational web survey, even though data from the first survey did not support the management's fear that employees would respond more than once (Church 2001). Although the possibility of mischievous multiple submissions cannot be ruled out, it seems unlikely that respondents would waste their time in this way (Reips 2002b), especially if there is no incentive to do so. A detailed analysis of eighteen repeat submissions from the four hundred complete responses to a survey of library users concluded that the majority resulted from innocent causes (Steffensen 2004), but the use of financial incentives may encourage dishonest repeat submissions (Batagelj & Vehovar 1998). Thus, whereas one web survey of gay and bisexual men excluded just 0.4% of responses on the basis of apparently intentional deception or deviousness (Ross *et al.* 2000), a similar survey which offered a $20 incentive found that 10% of its complete responses were repeat submissions, with one respondent submitting the survey 65 times (Konstan *et al.* 2005).

The final type of online non-probability sampling is that of ***volunteer panels of Internet users*** which Internet pollsters, YouGov, have championed in their UK public opinion polls. YouGov draws samples from its 46,000 panel of volunteers and weights responses according to various demographic and attitudinal variables in order to achieve its goal of "*modelling public opinion*" with speed and accuracy (Kellner 2004). It has been particularly successful in predicting election outcomes. Sparrow and Curtice (2004), however, found that responses given by online

respondents on various issues had no apparently discernable pattern when compared to responses from RDD samples. After demographically weighting the Internet sample according to the general population, some issues remained where the Internet sample did not reflect the general population sample. This led them to conclude that it was not possible to decide when an Internet poll was a reasonable estimate and when it was not and that the use of Internet polls for the type of untestable data that is collected in opinion polls was therefore risky (Baker *et al.* 2003). Kellner (2004) has continued to defend the use of weighted samples of Internet volunteers, arguing that it is not appropriate to judge online methods against the benchmark of traditional polling where there is no consensus or perfection. This chapter aims to contribute to the debate surrounding use of online volunteer samples, while focusing on a population that is difficult to reach through probability methods.

6.1.4.3 Who responds to Internet surveys?

Before turning to online research involving samples of gay and bisexual populations, the following will explore the composition of online samples that have been drawn from other populations. The Internet population has generally been described as predominantly young, male, educated and of higher socio-economic status, with respondents to Internet surveys likely to reflect this profile (Sheehan & Hoy 1999). This bias is not reflected in all Internet samples, however, with a survey of library users finding that the web and exit samples comprised similar library patrons (Perkins & Yuan 2001) and an email survey of physicians finding that the demographic characteristics of the respondents were representative of the population of physicians from which they were drawn (Alves & Szucs 2001). In another case, although male students were indeed less likely to respond to a postal survey than a web survey, there were no significant differences between samples in terms of age, ethnicity or academic-related variables (McCabe *et al.* 2002). This survey, which randomly assigned students to a web or postal mode, found that females were more likely to respond to both the postal and the web survey, which is evidence that certain propensities to respond may overcome modal differences.

Table 6.2.1: Previous studies comparing characteristics of non-Internet and Internet samples

Study	Non-Internet sample	Internet sample	Findings for Internet sample
Smith & Leigh (1997)	56 student volunteers from subject pool	72 volunteers from psychology research newsgroup	males over-represented; no differences in education or ethnicity
Zhang (1999)	31 researchers responding to survey request via post or fax	125 researchers responding to survey request via web	no differences in gender; younger
Bailey *et al.* (2000)	402 students attending selected classes	7,217 online volunteers	younger; wealthier; less ethnically diverse
Etter & Perneger (2001)	2,961 smoking cessation trial volunteers; 211 randomly selected smokers	1,027 volunteers from smoking cessation web site	males over-represented; younger; more education
Ballard & Prine (2002)	1,624 citizens responding to survey request via post	663 citizens responding to survey request via email	no differences in gender; younger; wealthier; more education; less ethnically diverse
Best & Krueger (2002)	710 voters randomly selected by telephone number	468 voters randomly selected from email listings	males over-represented; younger; wealthier
Kwak & Radler (2002)	402 students randomly selected from register	270 students randomly selected from register	males over-represented; younger
McDevitt & Small (2002)	210 triathlon participants who registered by post	132 triathlon participants who registered online	no differences in gender or age; wealthier; less education

Table 6.2.2: Previous studies comparing characteristics of non-Internet and Internet samples

Study	Non-Internet sample	Internet sample	Findings
Miller *et al.* (2002)	2,141 residents randomly selected from general population	774 residents randomly selected from general population with unknown Internet access	males over-represented; younger; wealthier
Bandilla *et al.* (2003)	1,485 residents randomly selected from general population	475 members of representative online panel	males over-represented; younger; more education
Chang & Krosnick (2003)	1,506 RDD-sampled adults	4,933 members of randomly selected online panel; 2,306 members of volunteer online panel	males over-represented; younger; wealthier; more education; white over-represented
Faas & Schoen (2004)	1,665 residents randomly selected from general population	598 members randomly selected from online panel; 29,853 online volunteers	males over-represented; younger; more education

The summary of studies comparing sample characteristics in Tables 6.2.1 and 6.2.2 shows, however, that an over-representation of males is frequently found in Internet samples (Bandilla *et al.* 2003; Best & Krueger 2002; Chang & Krosnick 2003; Etter & Perneger 2001; Faas & Schoen 2004; Kwak & Radler 2002; Miller *et al.* 2002; Smith & Leigh 1997), although gender is not always associated with mode of response (Ballard & Prine 2002; McDevitt & Small 2002; Zhang 1999). Internet samples tend to be younger (Bailey *et al.* 2000; Ballard & Prine 2002; Bandilla *et al.* 2003; Best & Krueger 2002; Chang & Krosnick 2003; Etter & Perneger 2001; Faas & Schoen 2004; Kwak & Radler 2002; Miller *et al.* 2002; Zhang 1999;) but age differences are not always found (McDevitt & Small 2002). They are also likely to be wealthier (Bailey *et al.* 2000; Ballard & Prine 2002; Best & Krueger 2002; Chang & Krosnick 2003; McDevitt & Small 2002; Miller *et al.* 2002). Findings with regard to education are more diverse. Internet samples have been found to be better educated (Ballard & Prine 2002; Bandilla *et al.* 2003; Chang & Krosnick 2003; Etter & Perneger 2001; Faas & Schoen 2004) but lower levels of education among Internet respondents have been reported (McDevitt & Small 2002) and differences are not always found (Smith & Leigh 1997). Findings on ethnicity are also less conclusive. Studies have reported that Internet samples under-represent black and over-represent white respondents (Chang & Krosnick 2003), are less ethnically diverse (Bailey *et al.* 2000; Ballard & Prine 2002) or that there are no differences in the ethnicity of online and offline samples (Smith & Leigh 1997). This may be due to the reluctance of ethnic minority respondents to participate in research in general, with sample comparisons finding that black respondents are under-represented in both modes (Bailey *et al.* 2000; McCabe *et al.* 2002).

The consistency reported in the above studies with regard to differences in gender, age and wealth is convincing. Most of these studies did not randomly assign respondents to the survey mode, however, as described in Tables 6.2.1 and 6.2.2, and caution is therefore advised in attributing causality to this aspect of the survey design. The studies also indicate that certain people may have a propensity or reluctance to respond whether by online or offline mode, with females more likely and ethnic minority individuals less likely to do so. In the same way, the citizens of

Thomasville, Georgia, had similar perceptions of community policing whether they returned the survey by post or by Internet (Ballard & Prine 2002), which the authors attributed to a consistent response bias across modes. Another citizen survey also found that responses to issues of community and service quality were consistent across web and postal surveys (Miller *et al.* 2002), which may result from a similar response bias.

Of particular relevance to the analysis undertaken here, is a comparison of a population based sample of respondents to the Swedish National Sexual Life Survey and respondents to a web survey that was advertised on banners in a Swedish sexual contact website (Ross *et al.* 2005). The Internet respondents were more likely to be young, highly educated, current students who lived in a city, identified as gay or bisexual and reported more than one sexual partner; they were less likely to be retired or in a current relationship. The present analysis will explore the extent to which these differences hold in a similar comparison of British gay and bisexual men.

It seems evident that differences between online and offline samples are likely to be found in Internet-related variables. Thus, nearly two thirds of telephone respondents to the third Slovene national web survey were non-intensive Internet users while only one quarter of web respondents were described in these terms (Vehovar *et al.* 1999). Other studies have also found Internet respondents to be keener Internet users (Kwak & Radler 2002; Zhang 1999), more technically knowledgeable (Kwak & Radler 2002) and with a higher perception of their ability to use the web (Zhang 1999). Along with differences in equipment capacity, such discrepancies in user ability have no doubt influenced the likelihood of responding to Internet surveys (Bradley 1999) although recent advances in both capabilities may now be bridging the divide (McDevitt & Small 2002).

As previously discussed, researchers have used weighting in an attempt to overcome the differences in online and offline samples. Although holding demographic background constant virtually eliminated the differences between web and pen-and-

paper responses to an environmental questionnaire (Bandilla *et al.* 2003), the use of sociodemographic variables to weight data has not generally proved to be satisfactory (Sparrow & Curtice 2004; Vehovar *et al.* 1999; Witte *et al.* 2000). A comparison of opinion polling by telephone and email also concluded that any representativeness of the broader public captured by the email survey was coincidental (Best & Krueger 2002). Email respondents were more likely to be male, younger, wealthier and more liberal and although weighting their responses improved the fit on some variables, it decreased the fit on others.

Although more complex weights have been successfully applied to Internet samples to derive population estimates of variables such as voting intention (Kellner 2004), the utility of such weights is limited because of their uncertain outcome when applied to other variables. Even when the demographic profile of online and offline respondents is similar, we cannot infer that these respondents will be identical on other variables of interest (Couper 2001) and if we cannot make this association, we cannot expect to derive weights that will be appropriate for the range of variables in which we may be interested.

Thus, the work undertaken here is not intended to improve the precision of our estimates of high risk sexual behaviour through suggesting how weights may be applied to Internet samples of gay and bisexual men. Its aim is to help shape our interpretation of the estimates derived from Internet samples by providing benchmark data on how such samples are likely to compare to probability samples on the variables investigated here.

6.1.5 Samples in surveys of sexual behaviour

It has been shown that the elderly and members of lower socio-economic groups are the most likely non-respondents to surveys in general (Goyder 1987). Given the sensitive nature of sexual content, it is important to consider the impact of this on respondents' likelihood of response and how it may influence the final composition of the sample. Any such differences are of particular relevance to this comparison of self-selected volunteers in the Internet and HIV web surveys and the randomly

selected respondents to Natsal who may not have participated in a survey of sexual behaviour on their own initiative. The following will discuss the issue in relation to surveys of sexual behaviour that have sought to generalise to a population of individuals with diverse sexualities, before examining the response of gay and bisexual men.

The research indicates that surveys of sexual behaviour are particularly prone to response bias. Although the overall response rate to Natsal 1990 was reasonably strong at 65% (Wadsworth *et al.* 1993), respondents were less likely to be male, older or living in inner city populations compared to census data, although minority ethnic representation was found to be comparable (Johnson & Copas 1997). Higher refusal of the self-administered section of the survey was associated with being from an ethnic minority, older, of lower occupational class, having comprehension difficulties, living in a metropolitan area and having one partner since the age of 13 (Copas *et al.* 1997). A Danish survey of sexual behaviour also found that males, older people and city dwellers were less likely to respond (Biggar & Melbye 1992). Although females, younger people and those who were more highly educated were also more likely to respond earlier, there was little difference between the lifestyles of these willing respondents and those who proved to be more reluctant. When respondents were randomly allocated to a sex survey or parent-child relationship survey, males were more likely to respond to the former and females to the latter (Barker & Perlman 1975). However, the overall response rates were similar for both surveys and as were respondent scores on a personality measure.

Other studies have shown how participation in surveys of sexual behaviour is related to the individual's sexual behaviour and attitudes. Participants in sex surveys are more likely to be novelty-seeking, eager to help, uninhibited personalities with liberal sexual attitudes and behaviours (Dunne *et al.* 1997) and such surveys are likely to attract people with the most interest in sex (Dunne 1998). It is argued that men who feel less positive about their homosexuality may be less willing to participate in sex surveys, whereas other samples of gay men may be more likely to participate from ideological motives (Sandfort 1997).

In comparison to nonrespondents, participants are more likely to have greater sexual experience (Catania *et al.* 1990; Strassberg & Lowe 1995) and to be more liberal in their sexual attitudes and report earlier sexual activity (Dunne *et al.* 1997). An analysis of those who refused to participate in an AIDS-related telephone survey and those who were difficult to contact revealed that these two groups of individuals were quite dissimilar (Turner 1999), concurring with earlier work indicating that the sociodemographic profile of different types of nonrespondent is quite different (Goyder 1987). Refusers were older, attended church more frequently, had less confidence in the confidentiality of surveys and felt greater discomfort about sexual self-disclosure. Those who were difficult to contact were more likely to be male, black, spend less time at home and have more sexual partners. Whereas the profile of refusers was associated with lower sexual risk behaviour, the profile of difficult-to-contact respondents was associated with higher risk behaviour.

Individuals who volunteer for surveys of sexual behaviour have sexual attitudes and behaviour that are consistent with the above findings. In comparison to students who volunteered for general social research, those who volunteered specifically for sex research had more sexual experience and a greater interest in sexual novelty, and scored more highly on nonconformity measures and sensation seeking (Bogaert 1996). From a group of students who participated in a mass testing session, those who volunteered for further studies of sexual behaviour reported less sexual guilt, more sexual experience and a more positive attitude towards sex, leading the authors to voice concerns over,

> *"the degree to which our exclusive reliance on volunteers ... may limit the populations to whom we might confidently apply our findings"* (Strassberg & Lowe 1995: 381-2).

The composition of the sample is likely to be affected by what led the respondents to participate in surveys of sexual behaviour. Thus, the profile of an Internet sample which mainly accessed a survey of sexual behaviour via a search engine using the terms "sex" and / or "survey", was different from the profile of college students who were invited to participate during classes in human sexuality or critical thinking

(Bailey *et al.* 2000). The greater sexual liberality and knowledge of the Internet sample cannot be divorced from the fact that they accessed the survey while actively pursuing an interest in sex. It has also been found that people who use the Internet for activities involving sexuality spend significantly longer online per week than those who do not (Cooper *et al.* 2002) which may give them increased opportunities to participate in web surveys of sexual behaviour.

These findings indicate that we should expect men who volunteered for the Internet and HIV web surveys to have greater sexual experience than those who were randomly selected into the Natsal sample. They suggest that the Internet samples which accessed the survey while online cruising will report higher levels of sexual activity.

6.1.6 Sampling minority populations

What follows is a discussion of sampling minority populations, with particular emphasis on gay and bisexual men. It begins with a description of the techniques that have been adopted to sample such populations and goes on to examine of how offline sampling approaches have shaped the composition of samples of gay and bisexual men. Given that the use of the Internet for social research is a relatively recent innovation, the subsequent description of online sampling will broaden to include other minority populations, before focusing on the composition of online samples of gay and bisexual men.

One of the problems that is common to all research involving minority populations is the low prevalence of potential respondents within the general population. Despite this, however, a number of probability techniques have been devised that may be suitable for sampling particular minority populations.

The first is ***screening from a general population sample***, which was mentioned earlier in relation to Natsal 2000's ethnic boost. This boost was screened by two procedures. The first involved screening all randomly selected households from postcode sectors where at least 12% of the resident population were from the ethnic

minorities of interest. The second involved a procedure called ***focused enumeration*** (Brown & Ritchie 1981) or ***multiplicity sampling*** (Kalton 1993; Kalton & Anderson 1986) whereby members of households from postcode sectors where at least 6% of the population were ethnic minority were also asked whether their neighbours were potentially eligible. The viability of this method depends on the prevalence of the minority population and may result in wasted effort if false positives who are not eligible for the study are identified and, more seriously, in sample bias if false negatives fail to be identified by the screening process (Lynn 1996b; Sudman & Kalton 1986). Although the low prevalence of gay and bisexual men may make screening difficult to justify on economical grounds, focused enumeration is untenable because it requires the population of interest to be a visible minority.

Natsal's ethnic boost also illustrates another technique for sampling minority populations, that of ***disproportionate stratification***. This is the division of the general population into strata representing different prevalence rates of the population of interest, in order to sample from strata with high concentrations of the minority population (Kalton 1993). In the case of Natsal 2000, postcode sectors were stratified according to the proportion of ethnic minority residents on the basis of data from the 1991 census. Although direct questions about sexual orientation were not included in the 2001 census, it might be possible to use proxy measures to locate high concentrations of gay and bisexual men. Thus, census data might be used to identify areas with higher prevalences of male couples and with higher proportions of single males. This approach has been successfully adopted in the United States, where areas have been selected according a range of criteria, including local informant information, residency of 35% or more never married adult males, reporting of at least 10 cases of AIDS, location of gay-related businesses and services and inclusion of residents on marketing lists targeting gay and bisexual men (Binson *et al.* 1996; Stall *et al.* 1996). The limitation of such techniques, however, is that only men living in areas with a higher prevalence of gay and bisexual men will be sampled.

When no adequate sampling frame exists for the minority population, it may be possible to generate one. This can be done using ***multiple frames*** which combine existing lists of members of the minority population (Kalton 1993; Kalton & Anderson 1986; Lynn 1996b). Membership lists from gay clubs or organisations might be used to create a sampling frame of gay and bisexual men. Although the frame is likely to contain missing elements, ineligible elements and duplicates, the procedure is useful in the absence of any alternative. Sampling frames of minority populations may also be created using ***snowballing***. Instead of stopping when the sample size is met, as described earlier, the procedure continues until no further new members of the minority population are located and the resulting list is used as the frame (Kalton 1993; Sudman & Kalton 1986). Such techniques can be readily applied to sampling gay and bisexual men, but their key limitation is that those who are socially isolated from other members of the population will be missing from the frame (Kalton & Anderson 1986).

It is also possible to capitalise on work undertaken in other studies by following up cases identified in ***previously collected probability samples*** or by adding screening questions to other surveys (Sudman & Kalton 1986). The lack of national probability surveys in the UK that have collected data on the sexual orientation means that the former is not an option for sampling gay and bisexual men in the UK, although it may be possible to add questions about sexual orientation to omnibus surveys in the future.

In the light of the problems described here, studies involving gay and bisexual men have generally adopted a convenience approach to sampling (Kalton 1993; Martin & Dean 1993; Rhodes *et al.* 2002). The Internet and HIV study used London gyms, HIV treatment clinics, HIV testing clinics and the Internet to recruit samples of gay and bisexual men. Other places commonly sampled include gay bars, organisations, events and so on. Although it is possible to use ***location sampling*** to generate a probability sample by choosing a location, such as a gay bar, then randomly selecting time periods and systematically sampling entrants to the bar, this strategy results in a

probability sample of visits to the selected venue (Kalton 1993) and does not help with the problem of generalisability to a population of gay and bisexual men.

6.1.6.1 Offline sampling of gay and bisexual men

The fact that less than 3% of British men in Natsal 2000 reported same sex activity in the previous 5 years (Mercer *et al.* 2004) means that the main challenge is to find an economic method for obtaining a sample (Kalton 1993). Given the low estimated prevalence of homosexuality in the general population, it is argued that,

> *"Few researchers have the time and financial resources to conduct surveys of sufficient size to gather a subsample of lesbian, gay male and bisexual respondents large enough to conduct rigorous analyses"* (Mathy *et al.* 2002: 253).

It is not surprising then that research involving the sexual behaviour of gay and bisexual men has generally relied on convenience samples obtained from clinic or community settings (Kalton 1993; Martin & Dean 1993; Rhodes *et al.* 2002), with one review finding that probability sampling was used in only 3 of 152 studies involving LGB samples (Sell & Petrulio 1996). At the same time, however, the findings from such research inform public health policy, service planning and resource allocation, making it of great importance to understand their generalisability (Martin & Dean 1993).

The only randomly selected samples of MSM across the UK consist of men drawn from the general population samples making up the two National Surveys of Sexual Attitudes and Lifestyles (1990 and 2000). The total number of men from the core sample who were sexually active with men in Natsal 2000 was 175 (before weighting), which is likely to limit the analytical potential of the sample. One solution to the small sample sizes attained by surveys of minority populations is to pool sample survey responses, but the absence of UK probability surveys of sexual behaviour other than Natsal means that this technique could not be adapted to estimate numbers of gay and bisexual men in the UK, let alone provide data for the analysis of their sexual behaviour. It was successfully used in the United States,

however, where the results of 56 Gallup surveys were combined to provide an estimate the number of Southern Jews which tallied with the estimation of the American Jewish Committee (Reed 1975).

Researchers in the United States have also undertaken successful probability surveys of urban gay men. Although such samples appear to under-represent segments of the population of gay and bisexual men because they are overly white, young and well-educated (Harry 1990), this may be an artefact of the urban migrants that they represent (Catania *et al.* 2001a). These surveys have also tended to randomly select from areas with high densities of gay and bisexual men who may not be representative of the population of MSM. The Urban Men's Health Study was a telephone survey of gay and bisexual men in four US cities which selected an RDD sample from areas with higher densities of the target population and screened responses for eligibility. 82% of these urban gay and bisexual men were in-migrants and their demographic profile was a reasonable reflection of such in-migrants who tend to be white, young and childless (Catania *et al.* 2001a). Another approach, which is similar to focused enumeration, is to screen members of an RDD sample, then use telephone numbers which located clusters of gay and bisexual men to generate further numerically similar telephone numbers (Blair 1999). These numbers are also screened for eligible respondents, again resulting in a sample of gay and bisexual men who are geographically clustered.

Although the San Francisco Men's Health Study also sampled randomly from an area of San Francisco with a high gay population, it stated that the findings should only be generalised to the population of single men living in the study area (Winkelstein *et al.* 1987). This conforms to the findings of a study by Mills *et al.* (2001) which found that MSM living in areas described as "gay ghettos" were different from those who did not and were more likely to be white, young, on high incomes, gay identified, open about their sexuality, involved in domestic partnerships in the gay community. By contrast, over half of the MSM identified by a national RDD sample did not have a regular gay friend, indicating their isolation from the gay community (Harry 1990).

These studies indicate that there may be a degree of commonality among certain sections of the male gay and bisexual population, but there are also major differences and the profile of the sample is likely to depend on how and where it was recruited (Sandfort 1997; Sell & Petrulio 1996; Weinberg 1970). Gay men sampled from bars are, for example, more likely to be single, young and sexually active than gay men in general (Sell & Petrulio 1996). It is important, therefore, to consider how the setting from which the respondents were recruited may shape the composition of the sample.

The work of Martin and Dean (1993) indicated that samples of men who are affiliated to gay groups and attend gay events are likely to be fairly similar whereas clinic samples may be somewhat different. Their non-probability sample of New York gay men was recruited from five sources in an attempt to draw a sample that was more representative of the diverse population of New York gay and bisexual men. It consisted of members of targeted gay organisations, unsolicited volunteers who had heard about the study, men referred by respondents from the pilot study, patients from STD clinics and men who attended a gay pride festival. These men sent information about the study to three other men, creating a snowball sample. The clinic sample were more likely to be young, less educated and on low incomes and less likely to be white or affliated to gay groups, but the differences between the other four groups were fairly small. This sample of New York men compared well as a whole with two probability samples of San Francisco men in terms of age and ethnicity, although it was also found to be more highly educated.

A Dutch study which compared a random sample of gay men with three convenience samples, however, concluded that the convenience sample estimates of both sociodemographic characteristics and sexual behaviour may be biased (Sandfort 1997). The random sample was found to have a lower level of education than men who completed a questionnaire placed in a gay magazine and those who were recruited via a gay and lesbian organisation. Members of the random sample who were in a steady relationship were also less likely to have had sexual contact with other men than the men with partners from the convenience samples, indicating that convenience samples may over-estimate sexual activity.

A small-scale comparison of 26 gay men recruited via a gay community source in New York with a sample of 52 New York gay men recruited through RDD, showed that the RDD sample were less affiliated to the gay community, were less likely to be open about their orientation and were more likely to have a regular sexual partner (Meyer & Colten 1999). Both these samples of gay men contained fewer minority ethnic men than were found in the sample of straight men who were recruited via RDD.

These findings from comparisons of convenience and probability samples suggest that we can expect to find differences between the Internet and HIV and Natsal samples both in terms of sociodemographic profile and sexual behaviour.

6.1.6.2 Online sampling of minority populations

A number of studies have pointed to the utility of the Internet for accessing minority populations (eg Rhodes *et al.* 2002) and the techniques described above for recruiting respondents into self-selected web surveys have been successfully adopted for sampling gay and bisexual men.

The Internet and HIV study placed pop-ups and banners in the two most popular websites for gay men in the UK. Over four-week periods in May to June, 2002 and 2003, samples of 4,500 and 2,500 UK men (respectively) were achieved. Sigma research has added a similar web survey to its national surveys of gay and bisexual men, which recruited seven and a half thousand men in 2002 and nearly eleven thousand men in 2003 (Hickson *et al.* 2003; Reid *et al.* 2004). Ross *et al.* (2000) placed their web survey of gay and bisexual men on the website of the Swedish Federation for Lesbian and Gay Rights and received a total of 678 responses over a six-month period. Rhodes *et al.* (2002) used a combination of methods to recruit non-probability samples of gay and bisexual men for their web survey. Having identified gay and bisexually orientated websites through online searches, they asked the webmasters of these sites to establish links to their data collection site, they posted the survey URL in online "guest books" and announced the survey in electronic mailing lists for gay and bisexual interest groups. Over a one-month

period, they received over 620 responses and achieved a sample of 382 gay and bisexual men living in the United States.

Although the size of samples generated via the Internet is no guarantee of data quality, one important advantage of large numbers is that it makes analysis of narrowly defined subgroups possible (Etter & Perneger 2001). In the same way, the Internet provides an effective means of recruiting reasonably sized samples of certain narrowly defined target populations. For example, a sample of 112 HIV positive barebackers (men who have intentional unprotected anal sex) was successfully recruited via four listservs and one chat room (Halkitis & Parsons 2003). Although it would be possible to reach such a sample by other means, it would probably take considerably longer (Wright 2005).

The Internet's facility to draw together geographically dispersed people who share common interests and the anonymity that may surround such encounters are likely to have contributed to its success in reaching minority populations. Koch and Schockman (1998) described the Internet as a "virtual homeland" bringing together a lesbian, gay and bisexual diaspora that is scattered around the globe. Use of the Internet to research minority populations also benefits from the Internet's capacity to create networks. Although researchers should not rely on search engines to secure participants (O'Neil 2002), an invitation to participate in a web survey that was posted on a listserv for involuntary celibates prompted webmasters from other relevant websites to create links to the survey and search engines to generate further links, resulting in a final sample of three hundred respondents (Burgess *et al.* 2001). The authors felt that the Internet successfully provided access to a geographically dispersed, stigmatised and difficult to specify population.

Some researchers have entered chat rooms and sought to recruit respondents by actively soliciting participation or passively waiting to be approached. Mathy *et al.* (2002) identified an exhaustive list of chat rooms targeted at lesbian and bisexual women using major search engines and approached randomly sampled potential respondents from the list of people present in these chat rooms, resulting in a sample

of 82 lesbians and bisexual women. Fernandez *et al.* (2004) spent over two hundred hours in chat rooms for Hispanic MSM, engaging over seven hundred chatters to generate a sample of 171 men. Although neither study reported any negative consequences to these approaches, another study of gay and bisexual men found that soliciting participation in chat rooms was not appreciated (Bull *et al.* 2001). The suggested alternative was to recruit by "lurking" in chat rooms and waiting to be approached. Those who were interested could chat to the researcher about the project and be directed to a website where the project details were held. A similar approach was taken in a study of Chinese gay and bisexual men (Wang & Ross 2002), where researchers lurked in the chat rooms of seven gay websites and waited for respondents to approach them. Despite this intense use of resources, only 53 useable questionnaires were returned, compared to 300 that were returned in response to emailing 452 users of the same websites.

Another method is to recruit respondents through offline community associations. A web survey of lesbian, gay, bisexual and transgender (LGBT) college students in the US recruited 450 participants by sending information on the survey to LGBT college organisations (Lindley *et al.* 2003). Although this approach relies on the good will of organisations to advertise the survey, it illustrates how such organisations provide access to community affiliated respondents whereas other minority populations, such as those involved in illegal activities, may not have such community affiliations. Coomber (1997), for example, was interested in the practices of drugs dealers. His respondents were recruited from 23 drug-related newsgroups, resulting in a sample of 80 individuals from across the world. He points out that it is not possible to define, let alone design, a probability sample of drug dealers and argues that, although the problems associated with online convenience sampling cannot be ignored, the data generated can lead research in new and exciting directions.

Duncan *et al.* (2003) also found the Internet to be an effective means of researching illegal activities in their research involving to nonabusive drug users. The registration of the web survey site with major search engines and postings to electronic mailing lists and USENET news groups, resulted in a sample of nearly

2000 people by the third wave, which the authors describe as the largest sample of nonabusing illicit drug users ever assembled. It may be, however, that the apparent increase in the prevalence of recreational drug use may make the use of probability sampling more feasible in such research, with the British Crime Survey 2000 finding that 34% of adults had used drugs in their lifetime, rising to 58% of 20 to 24 year olds (Ramsay *et al.* 2001).

6.1.6.3 Who responds to Internet surveys of gay and bisexual men?

It is not surprising that studies have indicated that sampling gay and bisexual men from the Internet may succeed in reaching men who are less acculturated into the gay community, given that most of the comparison samples were recruited through their community associations. Thus, a comparison of respondents to a web survey hosted by a gay website with respondents to a pen-and-paper survey distributed through gay community contacts found that Internet respondents were younger, more likely to live in small towns, live with their parents or a girlfriend and have a lower education (Ross *et al.* 2000). They were also more likely to be behaviourally bisexual and to self-identify as bisexual but the sexual practices of the two samples were generally similar. It was also found that the men who used gay chat rooms were more likely to be young, self-identify as bisexual, live outside major cities, report UAI and not to have tested for HIV than those who did not (Tikkanen & Ross 2000), again indicating that web surveys may access respondents with a different profile to those reached via offline gay community recruitment.

Similarly, when Rhodes *et al.* (2002) compared samples of gay men recruited from gay and bisexually orientated websites and mailing lists with men recruited in gay bars, the Internet sample was more likely to identify as bisexual than gay. This Internet sample was older than the community sample, more likely to be HIV positive and have a history of STDs. Although these findings may indicate higher risk behaviour among the Internet sample, respondents were not randomly assigned to survey mode and the differences might also be attributed to respondent hesitation over disclosing bisexual orientation and STDs in the context of the bar (Rhodes *et al.* 2002).

Comparison of the 2002 Internet sample of the Internet and HIV study with the 2002 gym sample found that Internet respondents were younger and less educated (Elford *et al.* 2004b). They were less likely to have sex only with men, be in a relationship or to have tested for HIV. Further analysis showed that HIV positive Internet respondents were broadly similar to those surveyed in the gyms, whereas HIV negative Internet respondents and those who had never tested for HIV were less likely to self-identify as gay, more likely to be behaviourally bisexual and more likely to report high-risk sexual behaviour than their offline counterparts.

Sigma research compared samples from its 2002 survey drawn from three sources (Hickson *et al.* 2003). The pride sample was recruited at seven national gay pride events, the booklet sample was recruited through questionnaires distributed by agencies mainly involved in health promotion and the Internet sample responded to pop-ups and banners in gay.com. In comparison to the other two samples, the Internet respondents were more likely to be young, white, self-reported middle class and behaviourally bisexual. They were less likely to identify as gay, to have tested for HIV or to be HIV positive. Additionally, the 2003 booklet sample were more likely to have had larger numbers of sexual partners than the 2003 Internet sample which the authors highlight as contrary to the hype around high Internet sexual activity (Reid *et al.* 2004).

Rhodes *et al.* (2003) argue that the higher prevalence of lower educated persons that has been found among some online populations of gay and bisexual men may reflect an adoption of the Internet by some marginalized groups, resulting in a more diverse sociodemographic profile among Internet samples of gay and bisexual men than among samples of the general population of Internet users. Koch and Schockman (1998) also argue that the Internet facilitates access to a diverse population of lesbians, gay men and bisexuals. They collected information on the seven thousand people who registered with the site which was set up to disseminate information about the "Queer Cyber Center". Although the largest group represented were young, white, college-educated gay men, the registrants were diverse, indicating that the Internet had the capacity to reach a broad spectrum of LGB men and women.

As previously discussed, it is important to consider how and where the sample was recruited and the effect that this may have on its composition. The Internet is not a homogenous location but a collection of diverse virtual settings and in the same way that sampling from particular offline settings has been shown the influence the composition of the sample, the selection of particular online settings is likely to have a similar influence. This diversity undoubtedly extends to use of the Internet for sexual purposes and, as Ross *et al.* have argued,

> *"it would be a mistake to characterize the Internet sexual subculture as monolithic"* (Ross *et al.* 2000: 752).

In this way, a study that recruited 1,776 men into an online STD/HIV prevention, found that different websites for meeting same sex sexual partners attracted different men (Bull *et al.* 2004). America Online was more frequently named by men who were older, whereas gay.com was more popular among younger men and MSM (rather than men who had sex with men and women or MSMW), and Yahoo was more commonly cited by MSMW and men with no college education.

Thus, the composition of the sample may be more affected by the websites targeted than the electronic method of communication with respondents. For example, few differences were found between Chinese gay and bisexual men whose survey participation was sought in seven gay chat rooms and those whose participation was solicited via email addresses from websites associated with the same seven chat rooms (Wang and Ross 2002).

The above review began with a description of probability and non-probability sampling and went on to describe how these methods have been applied to online sampling and sampling minority populations. It has sought to describe how the adoption of particular sampling strategies and the sexual content of the survey may impact on sample composition. The review concluded with a summary of the small number of studies which have examined the composition of Internet samples of gay and bisexual men. The following analysis seeks to add to this body of knowledge

through a comparison of the characteristics of self-selected Internet samples of gay and bisexual men with a sample of MSM drawn from a probability sample.

6.2 Methods

The Internet and HIV samples that will be examined here are the two Internet samples of gay and bisexual men who were recruited into the Internet and HIV web surveys in 2002 and 2003. The randomly selected sample consists of the MSM who participated in the National Survey of Sexual Attitudes and Lifestyles 2000. What follows is a summary of the points of importance in the design of the two studies for the secondary analysis undertaken here. A detailed explanation of the methods used in the Internet and HIV study can be found in the paper "*The Internet and HIV study: design and methods*" by Elford *et al.* (2004a) and the methods used by Natsal 2000 are described in "*National Survey of Sexual Attitudes and Lifestyles II: Technical Report*" by Erens *et al.* (2001).

6.2.1 Description of samples

Internet and HIV: Internet samples

The two Internet samples were recruited from gaydar (http://www.gaydar.co.uk/) and gay.com (http://uk.gay.com/), two of the most popular websites for gay men in the UK. These samples are referred to as the "Internet samples". Over a five-week period in May to June, 2002 and 2003 pop-ups and banners appeared in chatrooms and profile pages asking men to participate in the survey. Clicking on a pop-up or banner took respondents to the web survey which they could complete and submit online. Only respondents who said they were at least 18 years old were permitted to answer the questionnaire. Following questions on sociodemographic profile, including self-identified sexual orientation, respondents were asked the gender of their sexual partners in the past year. Only men who had had sex with a man in the past year were asked to continue. Although the pop-ups and banners were restricted to UK chatrooms or profiles, anyone entering these online locations from anywhere in the world over the survey period had the opportunity to participate.

Natsal sample

The aim of Natsal 2000 was to interview a representative sample of men and women aged 16 to 44 living in private households in Great Britain. The sample design is described in detail under *6.1.2 Probability sampling* and *6.1.6 Sampling minority*

populations and resulted in a core sample of the general population and a boost sample of black and Asian adults. Only MSM respondents selected from the core sample are included here. This sample is referred to as the "Natsal sample". Fieldwork took place from May 1999 to February 2001, with respondents interviewed by trained interviewers in their homes. The interview began and ended with face-to-face interviewing and a self-completion component was administered by CASI in the middle of the questionnaire. This component covered what were considered to be the most sensitive questions on sexual behaviour including sexual health and recreational drug use.

6.2.2 Sample comparison

Table 6.3 shows the periods of data collection in chronological order for the three samples used in this secondary analysis.

Table 6.3: Data collection for Natsal and Internet samples

Sample	Data collection period
Natsal 2000	May 1999 to February 2001
Internet 2002	May to June 2002
Internet 2003	May to June 2003

Although data were collected from the Internet 2003 sample two or more years after fieldwork ended for Natsal 2000, this sample is included in order to provide findings that can be understood in the context of the findings from the other chapters. It also contains more variables than the Internet 2002 sample that can be used to compare respondent characteristics.

The analysis which follows begins with a comparison of the Natsal sample with the 2003 and 2002 Internet samples at the national level. The focus then turns to London men and a comparison of the Natsal London sample with the 2003 and 2002 Internet

London samples, followed by a similar comparison for samples from outside London. The analysis concludes with a comparison of the Internet national samples recruited from gaydar and from gay.com.

There was a lack of consistency between the eligibility criteria for respondents to the different surveys and, in order to ensure the closest correspondence between members of each sample, respondents were only included in any of the above sample comparisons if they fulfilled the following criteria:

- were coded as male by the Natsal interviewer or completed the Internet and HIV web surveys that were explicitly for men;
- stated they were aged between 18 and 44;
- stated they had had at least one homosexual partner in the last year.

Men were included in the London sample comparisons if they:

- were coded as living in London for Natsal or stated they were currently living in London for the Internet and HIV web surveys.

Whereas the Internet samples comprise mainly men who self-identify as gay or bisexual (only 1% of UK men who completed the 2003 web survey identified as "straight/heterosexual" or "other"), Natsal did not contain any questions about sexual orientation. The self-identification of the Natsal men is therefore unknown.

The Natsal national sample contained men living in England, Scotland and Wales (Great Britain) whereas the Internet national samples also included men living in Northern Ireland (United Kingdom). Men living in Northern Ireland were therefore excluded from the 2003 Internet national sample for the purposes of this comparison. The 2002 Internet and HIV web survey did not establish where in the UK respondents lived and the 2002 Internet national sample may therefore include men living in Northern Ireland.

6.2.3 Sociodemographic and behavioural comparisons

Comparisons were made between the samples listed above on the following measures:

Sociodemographic

Age

Age group

Ethnicity

Work status

Highest educational qualification

Number of post-16 years in education

Lives in urban or rural area

Area of Great Britain where lives

Social class

Place of birth

Health status

Self-assessed health status

Whether tested for HIV

Alcohol and drug use

Frequency of drinking alcohol in last year

Whether injected drugs in last year

Sexual health and behaviour

Age of first sex with a male

Whether had STDs in last year

Whether had gonorrhoea in last year

Whether had syphilis in last year

Whether had another STD in last year

Whether has had recent anal sex

These variables were selected on the basis that they were included in the Natsal questionnaire and could be mapped onto responses from at least one of the Internet and HIV web surveys. In other words, there are some measures where only one of the Internet samples is compared to the Natsal sample. Appendix 5 provides a detailed explanation of how all these measures were derived from questions contained in the three questionnaires. Most of these measures are considered to be reasonably commensurate across the surveys and to operationalise the same concepts. Although it is acknowledged that small differences in question wording and format may influence response patterns (Schuman & Presser 1981; Smith 1979), the following will identify differences in questions that are considered to be of particular concern, where underlying concepts are somewhat different or operationalisations are manifestly distinct. It will explain how such differences were

addressed in the analysis where this was possible and suggest how they may have affected the comparability of the measures where it was not.

Social class was coded using the National Statistics Socio-economic Classification (NS-SEC) and translated into social class groups I to V. It should be noted that coding social class is a complex matter involving a number of variables and extensive lists of tables and that the reliability of the coding is likely to vary with the experience of the coder (Elliot 1982). Furthermore, the data that were available for coding in Natsal and the Internet and HIV study were slightly different. Thus, Natsal respondents who were students but also had work experience were coded for social class on the basis of previous work experience whereas Internet respondents who were students were not asked about previous employment and could not be coded for social class. For this reason, Natsal respondents who were currently students were excluded from the social class comparisons.

The ***number of post-16 years in education*** is calculated from a Natsal question asking at what age respondents completed their full-time *continuous* education whereas Internet respondents were asked to report their total number of post-16 years in full-time education. Internet respondents may therefore include non-continuous full-time education in their calculation and, other things being equal, may report more years of post-16 education.

The measurement of ***self-assessed health status*** was somewhat different. Natsal asked respondents about their health in general whereas the Internet and HIV surveys asked about health in the past 3 months. The scales adopted may also have influenced response, with "fair" representing the Natsal mid-point and "good" representing the Internet and HIV mid-point, as follows:

Natsal	Very good	Good	Fair	Bad	Very bad
Internet and HIV	Excellent	Very good	Good	Fair	Poor

New scale	Very good	Good	Fair	Poor

The above new four point scale was created to integrate the data from Natsal and the Internet and HIV web surveys. Schwarz *et al.* (1985) found that respondents used the categories available to deduce a range of expected behaviours within which to locate themselves. When the categories were combined to create the new scale, other things being equal, the Internet respondents may report better health because the options available to them emphasise the positive end of the scale (Tourangeau *et al.* 2000).

Although Internet respondents were asked directly if they had ***had any STDs in the last year,*** Natsal respondents were asked the year that they were last diagnosed with STD "x" and not the exact date. If, for example, they reported diagnosis in 1999 and their year of interview was 2000, this diagnosis was coded as "in the last year", although it could have been up to 2 years ago. Other things being equal, Natsal respondents may report more STDs because their response may refer to a longer time period.

In asking whether respondents had ***tested for HIV,*** Natsal respondents were given the option to say that they were "not sure" whereas Internet respondents had to select from the "yes"/"no" options available. It is argued that where a "not sure" option is provided, respondents may satisfice or fail to expend the necessary effort to generate an optimal answer (Krosnick 1991; Krosnick & Fabrigar 1997). 11 of the Natsal respondents selected this option and since we cannot infer how they would have responded if they were forced to choose between "yes" and "no", these respondents were excluded from the HIV testing comparisons.

When asked whether they ***injected drugs in the last year,*** Natsal respondents were asked about any non-prescription drugs that they had injected whereas Internet respondents were asked about any drugs except steroids that they had injected which may include prescription drugs. Internet respondents were asked about drug injection following a question about recreational drug use and this context is likely to have indicated that the question was about injecting non-prescription drugs (Tourangeau *et al.* 2000). A degree of ambiguity remains, however, as to whether

the question includes injection of prescription drugs. Thus, other things being equal, Natsal respondents may report more drug injecting if they include non-prescription steroids and Internet respondents may report more drug injecting if they include any prescription drugs.

The scales adopted for ***frequency of drinking alcohol in the last year*** were somewhat different, with the mid-point on the Internet and HIV scale equal to a lower frequency of drinking than the mid-point on the Natsal scale (see Appendix 5). Schwarz *et al.* (1985) found that respondents reported less television viewing when the scale ranged from "up to a half hour" to "more than two and a half hours" than when it ranged from "up to two and a half hours" to "more than four and a half hours". Other things being equal, Internet respondents may similarly report less frequent drinking than the Natsal respondents.

In determining the respondent's ***age of first sex with a male***, Natsal asked one question about any kind of sexual experience or sexual contact and another question specifically about genital contact. Answers to both of these questions are used to compare with the Internet and HIV question which asks about "sex of any kind", in accordance with the view that,

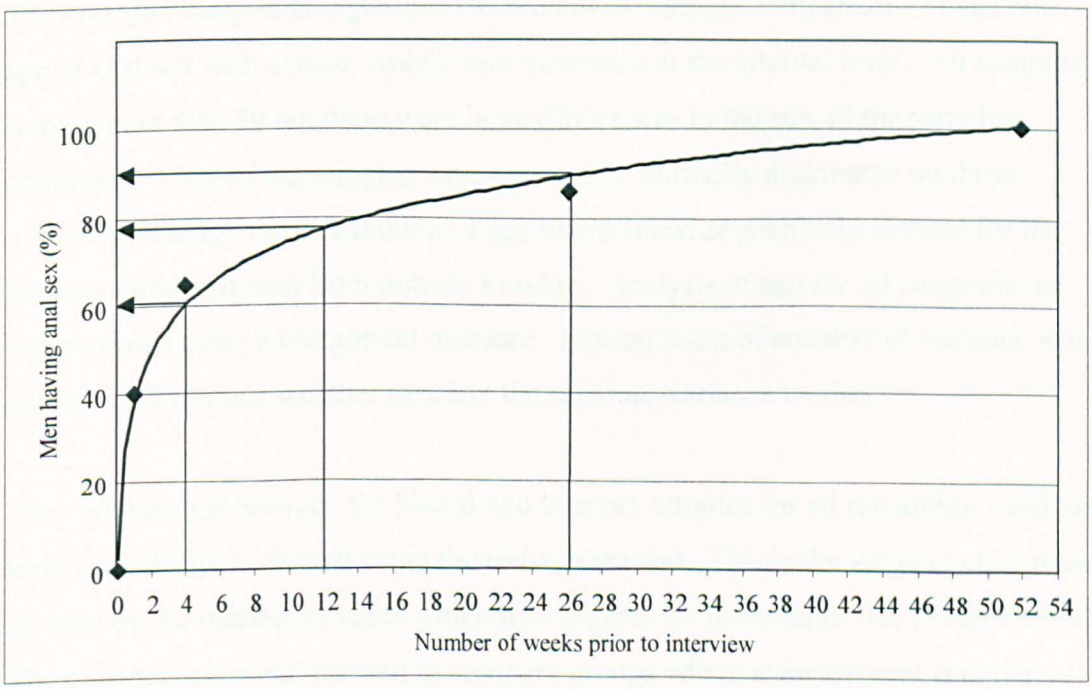
> *"there is no consensus on an act that defines first homosexual experience"* (Michaels & Giami 1999: 417).

Internet respondents were asked if they had had anal sex in the previous three months, whereas Natsal respondents are asked whether they had had anal sex in the previous week, 4 weeks, 6 months or year. In this way, measures of ***recent anal sex*** do not correspond. Other things being equal, Internet respondents should report more anal sex in the past 3 months than reported by Natsal respondents in the past 4 weeks and they should report less anal sex in the past 3 months than reported by Natsal respondents in the past 6 months.

Figure 6.1 (on the following page) is a plot of the cumulative percentage of Natsal respondents who reported having had anal sex in the period that was one week, four

weeks, six months and one year prior to interview. A logarithmic trend line is fitted to the Natsal estimates. It illustrates how the majority of men who have had anal sex in the past year are likely to report its recent occurrence and we would therefore expect a Natsal guesstimate at 3 months to be somewhat higher than the 4 week estimate and closer to the 6 month estimate, as indicated by the arrows on the diagram.

Figure 6.1: Cumulative percentage of Natsal men reporting anal sex in the past year



6.2.4 Data analysis

In general terms, weighting is used in probability samples to correct for unequal probabilities of selection, to adjust for non-response and for post-stratification in order to match the sample to the population profile (Kalton 1983; Lynn 1996a). In both the general population sample and the boost sample, Natsal 2000 applied three sets of weights to correct for unequal probabilities of selection. The first corrected for the selection according to postcode sector, the second corrected for unequal selection of households at multiple occupancy addresses and the third corrected for

varying probabilities in selection of adults within households (see Erens *et al.* 2001). Weights were also applied to correct for non-response in the general population sample but not in the ethnic boost, because there were no reliable data to estimate non-response for the ethnic groups. Only data from the general population sample was used in the following comparisons and the appropriate weights were applied to the data prior to the following secondary analysis. No weights were applied to the Internet and HIV data because probabilities of selection and level of non-response cannot be calculated for these convenience samples.

The ***t test for independent groups*** was used in the sample comparisons of age and age of first sex with a male, which were measured at the interval level. All samples were greater than 30 but there were large differences in the size of the samples compared. Most of the samples were reasonably normally distributed on these variables although the distribution of age was somewhat positively skewed for the Internet sample of men from outside London. Analysis of age for all comparisons was repeated using a categorical measure. Levene's test of equality of variance was applied to determine whether to select the separate-variance t value.

The comparisons between the Natsal and Internet samples for all remaining variables were statistically evaluated using the ***phi-square test***. This is the value of chi-square divided by the number of cases which thus adjusts for the sample size (Everitt 1992). The chi-square test can be used to compare groups where measurement is at the nominal level. It examines the difference between observed frequencies and expected frequencies. The Natsal sample was considerably smaller than the Internet samples, however, with 171 men (before weighting) in the Natsal national sample used here, compared to Internet national samples of two to four thousand for the purposes of the following comparisons. In order to rule out the impact of these differences, the size of the Natsal sample was scaled up to the size of the Internet sample with which it was to be compared. The proportion of Natsal respondents in each category of the variable under comparison was applied to the total number of respondents in the Internet sample for that variable comparison and a new set of frequencies was derived for the Natsal sample.

Table 6.4: Derivation of scaled up Natsal frequencies

Age group	Original frequency	Percentage	Scaled up frequency (x 2065)
18 to 19	11	1.1	22.7
20 to 24	27	12.4	256.1
25 to 34	44	51.2	1057.3
35 to 44	35	35.3	729.0
Total	*117*	*100.0*	*2065.0*

Table 6.4 shows how a new set of frequencies for the national Natsal sample was calculated for the comparison of age group. The first column shows the original frequencies for the 117 Natsal respondents (after weighting) involved in the comparison. The second column shows these frequencies as percentages. The third column shows the scaled up frequencies which result from multiplying the percentages by 2,065 which is the number of Internet 2003 respondents with which the Natsal sample was compared on age group.

The new scaled up frequencies were then used as the expected frequencies (f_e) in a single sample chi-square test to calculate the value of chi-square (χ^2_{obt}) where the equivalent Internet and HIV frequencies were used as the observed frequencies (f_o):

$$\chi^2_{obt} = \Sigma \frac{(f_o - f_e)}{f_e}$$

Because of the large sample sizes used to calculate chi-square, the observed frequencies were found to be significantly different from the expected frequencies in the vast majority of cases which helped little with the interpretation of the findings. Thus, phi-square was calculated by dividing χ^2_{obt} by the sample size, which in the above example would be 2,065. Phi-square is a measure of "goodness of fit" such that a value of zero indicates a good fit between the Natsal and Internet sample with

which it is being compared and a value of one indicates a poor fit. It is the simplest way of overcoming the problem of the sample size although its interpretation is complicated by the fact that it may exceed an upper limit of one (Everitt 1992). The value of phi-square is reported here, along with the percentage differences between the samples, as an indication of the extent to which the distribution of the samples compared was statistically similar.

Chi-square and phi-square give an overall assessment of the difference between the observed and expected frequencies. Adjusted residuals can be calculated to isolate the cells within the contingency table which are responsible for this overall difference (Everitt 1992). These residuals represent the standardised differences between the observed and expected values for each cell which are then adjusted to account for the fact that the resulting variances are always less than or equal to one. The adjusted residuals were calculated for all comparisons but are not reported because they reflected the percentage differences between the two samples described in the following tables and did not add anything to the interpretation.

The ***chi-square test*** and the ***Mann-Whitney U test*** were also used to examine significant differences between respondents who started the 2003 web survey from gaydar and those who came from gay.com. It should be noted that data on website of origin were not available for 23% of UK respondents.

6.3 Findings

The following section begins with a comparison of the differences between the Natsal national sample of MSM and Internet national samples of gay and bisexual men from 2003 and 2002. This is followed by a comparison of the London men from the Natsal and Internet samples. The London comparison is included because of the relatively high number of gay and bisexual men in London compared to elsewhere in the United Kingdom, making this a key area for research among this population. Data collected from samples of men living outside London but within the British Isles are then presented. Finally, we return to the national samples to make a comparison of the differences between the Internet 2003 national sample of men recruited from gaydar and those recruited from gay.com.

6.3.1 Comparison of national samples

Comparisons between the national sample from Natsal 2000 and the 2003 and 2002 Internet national samples are shown in Tables 6.5.1 to 6.5.9 (on the following pages).

The Internet national samples were significantly younger than the Natsal national sample (mean ages 31.9 (s.d. 6.2) for Natsal vs 29.4 (s.d. 7.6) for Internet 2003 and 30.0 (s.d. 7.3) for Internet 2002, $p \leq 0.01$). This was reflected in the comparison of age groups which showed that only 1.1% of the Natsal respondents fell into the youngest age group (18 to 19) compared to 10.0% of Internet 2003 respondents and 8.1% of Internet 2002 respondents ($\Phi^2 = 0.87$ and 0.51 respectively). The Internet samples were statistically similar to the Natsal sample in terms of ethnicity ($\Phi^2 = 0.03$ and 0.02), with white respondents comprising 92.3% of the Natsal sample, 95.6% of the Internet 2003 sample and 95.0% of the Internet 2002 sample.

The work status of the Internet 2003 respondents was rather different from the Natsal respondents ($\Phi^2 = 0.31$). 17.0% of the Internet 2003 respondents were students compared to 5.2% of Natsal respondents. The Internet 2002 sample was a much better match for the Natsal sample ($\Phi^2 = 0.07$).

Table 6.5.1: Comparison of national samples

	Natsal	Internet 2003	Internet 2002
Mean age (yrs)	31.9	29.4	30.0
Standard deviation	6.2	7.6	7.3
Base (n)	*117*	*2065*	*3958*
t statistic		4.1**	3.1**

Note: *p ≤ 0.05, **p ≤ 0.01

Age group (%)

18 to 19	1.1	10.0	8.1
20 to 24	12.4	23.4	19.3
25 to 34	51.2	36.8	41.9
35 to 44	35.3	29.8	30.6
Base (n)	*123*	*2065*	*3958*
phi-square		0.87	0.51

Ethnicity (%)

White	92.3	95.6	95.0
Black African	0.3	0.1	
Black Caribbean	2.6	0.3	0.8
Other Black	0.0	0.1	
Asian	1.9	0.9	1.1
Mixed / Other	2.9	3.0	3.1
Base (n)	*117*	*2045*	*3958*
phi-square		0.03	0.02

Note: black respondents are collapsed into one category to calculate phi-square

Table 6.5.2: Comparison of national samples

	Natsal	Internet 2003	Internet 2002
Work status (%)			
Working	81.1	76.2	78.3
Not working	13.7	6.8	10.6
Student	5.2	17.0	11.1
Base (n)	*117*	*2065*	*3958*
phi-square		0.31	0.07
Highest educational qualification (%)			
Degree level	43.6	43.6	-
Not degree level	56.4	56.4	-
Base (n)	*117*	*2061*	-
phi-square		0.00	-
Number of post-16 years in education (%)			
None	21.8	14.5	16.1
Up to 2 years	20.4	23.4	22.0
3 years or more	54.1	49.4	50.8
Still in f-t education	3.7	12.6	11.1
Base (n)	*117*	*2056*	*3958*
phi-square		0.25	0.17

Table 6.5.3: Comparison of national samples

	Natsal	Internet 2003	Internet 2002
Lives in urban or rural area (%)			
Urban	93.7	91.9	90.2
Rural	6.3	8.1	9.8
Base (n)	*117*	*2046*	*3958*
phi-square		0.01	0.02
Area of Great Britain where lives (%)			
London	32.3	24.4	28.1
North East	5.0	2.9	
North West	11.6	12.5	
Yorkshire & Humber	8.4	6.4	
East Midlands	6.6	4.9	
West Midlands	6.5	8.1	
South West	2.8	6.9	
East of England	5.6	6.8	
South East	11.1	13.1	
Wales	5.8	3.4	
Scotland	4.5	10.5	
Outside London			71.9
Base (n)	*117*	*2018*	*3958*
phi-square		0.17	0.01

Table 6.5.4: Comparison of national samples

	Natsal	Internet 2003	Internet 2002
Social class (%)			
I	12.3	8.1	-
II	54.6	50.7	-
IIIN	13.2	23.5	-
IIIM	11.2	11.8	-
IV	7.6	5.0	-
V	1.1	0.9	-
Base (n)	*109*	*1577*	-
phi-square		0.11	-
Born in the United Kingdom or Eire (%)			
Born in UK / Eire	84.1	89.9	-
Not born in UK / Eire	15.9	10.1	-
Base (n)	*117*	*2054*	-
phi-square		0.03	-
Born in the European Union (%)			
Born in EU	89.5	93.6	-
Not born in EU	10.5	6.4	-
Base (n)	*117*	*2054*	-
phi-square		0.02	-

Table 6.5.5: Comparison of national samples

	Natsal	Internet 2003	Internet 2002
Global region of birth (%)			
Europe – EU	89.5	93.6	-
Europe – non-EU	1.3	0.8	-
Australia, New Zealand	1.9	0.8	-
North America	0.0	0.7	-
South & Central America	1.7	0.4	-
Caribbean countries	2.0	0.1	-
India, Pakistan, Bangladesh	1.1	0.1	-
China (inc Hong Kong, Taiwan)	0.0	0.6	-
Other Asia	0.5	0.9	-
Middle East, North Africa	0.2	0.3	-
West Africa	0.8	0.0	-
Central and East Africa	0.0	0.7	-
South Africa (republic of)	1.0	0.8	-
Base (n)	*117*	*2054*	-
phi-square		0.02	-
Self-assessed health status (%)			
Poor	2.9	3.6	-
Fair	8.6	16.3	-
Good	36.3	28.4	-
Very good	52.2	51.7	-
Base (n)	*123*	*2061*	-
phi-square		0.09	-

Note: categories are collapsed into "Europe", "Australia, NZ", "America and Caribbean", "Asia" and "Middle East and Africa" to calculate phi-square

Table 6.5.6: Comparison of national samples

	Natsal	Internet 2003	Internet 2002
Whether tested for HIV (%)			
Yes	52.9 (53.3)	52.8	53.4
No	46.4 (46.7)	45.8	46.6
Rather not say	0.6 (-)	1.4	-
Base (n)	*108*	*2065*	*3958*
phi-square		0.01	0.00
Note: "rather not say" category removed to calculate phi in comparison of Natsal and Internet 2002			
Frequency of drinking alcohol in last year (%)			
1 or 2 times a week	74.2	69.4	-
1 or 2 times a month	18.0	18.8	-
1 or 2 times a year	5.3	4.8	-
Never	2.5	7.1	-
Base (n)	123	2057	-
phi-square		0.09	-
Whether injected drugs in last year (%)			
No	99.1	98.2	-
Yes	0.9	1.8	-
Base (n)	*117*	*2056*	-
phi-square		0.01	-

Table 6.5.7: Comparison of national samples

	Natsal	Internet 2003	Internet 2002
Mean age of first sex with a male			
Age in years	17.9	17.6	-
Standard deviation	6.1	5.4	-
Base (n)	*117*	*2060*	-
t statistic		0.5	-

Note: comparison of "sex of any kind" (Internet and HIV) with "any sexual experience or contact (Natsal); *p ≤ 0.05, **p ≤ 0.01

Mean age of first (genital) sex with a male			
Age in years	18.1	17.6	-
Standard deviation	6.1	5.4	-
Base (n)	*115*	*2060*	-
t statistic		0.9	-

Note: comparison of "sex of any kind" (Internet and HIV) with "sex involving genital contact (Natsal); *p ≤ 0.05, **p ≤ 0.01

Whether had STDs in last year (%)			
Yes	4.8	16.9	19.0
No	95.2	83.1	81.0
Base (n)	*114*	*2053*	*3958*
phi-square		0.32	0.44

Table 6.5.8: Comparison of national samples

	Natsal	Internet 2003	Internet 2002
Whether had gonorrhoea in last year (%)			
Yes	1.3	-	5.4
No	98.7	-	94.6
Base (n)	*114*	-	*3958*
phi-square		-	0.13
Whether had syphilis in last year (%)			
Yes	0.0	-	1.3
No	100.0	-	98.7
Base (n)	*114*	-	*3958*
phi-square		-	-
Whether had another STD in last year (%)			
Yes	4.2	-	14.4
No	95.8	-	85.6
Base (n)	*114*	-	*3958*
phi-square		-	0.01

Note: phi cannot be calculated because no cases were expected to have had syphilis in the last year

Table 6.5.9: Comparison of national samples

	Natsal	Internet 2003	Internet 2002
Whether has had recent anal sex (4 weeks vs 3 months) (%)			
	4 weeks	3 months	
Yes	51.8	76.9	76.0
No	48.2	23.1	24.0
Base (n)	*117*	*2065*	*3958*
phi-square		0.25	0.23

Note: comparison of anal sex in the past 4 weeks (Natsal) with the past 3 months (Internet and HIV)

Whether has had recent anal sex (6 months vs 3 months) (%)			
	6 months	3 months	
Yes	69.3	76.9	76.0
No	30.7	23.1	24.0
Base (n)	*117*	*2065*	*3958*
phi-square		0.03	0.02

Note: comparison of anal sex in the past 6 months (Natsal) with the past 3 months (Internet and HIV)

Although Natsal and Internet 2003 national respondents were equally likely to be educated to degree level ($\Phi^2 = 0.00$), the samples were not so well matched in terms of the number of years that they had spent in post-16 education ($\Phi^2 = 0.25$ and 0.17). Both Internet samples were more likely to still be in full time education (3.7% for Natsal vs 12.6% for Internet 2003 and 11.1% for Internet 2002) and less likely to have had no post-16 years in education (21.8% for Natsal vs 14.5% for Internet 2003 and 16.1% for Internet 2002).

The division between urban-dwellers and rural-dwellers was very similar for the both Internet samples and the Natsal sample ($\Phi^2 = 0.01$ and 0.02) and the division between those living inside and outside London was very similar for the Internet 2002 and Natsal samples ($\Phi^2 = 0.01$). The Internet 2003 sample was a reasonably good match for the Natsal sample in terms of geographical distribution around the country ($\Phi^2 = 0.17$), although Natsal respondents were rather more likely to live in London (32.3% vs 24.4%) and Internet 2003 respondents were rather more likely to live in Scotland (10.5% vs 4.5%). The Internet 2003 sample was very similar to the Natsal sample on social class ($\Phi^2 = 0.11$). The two samples were also well matched on whether they were born in the United Kingdom or Eire ($\Phi^2 = 0.03$), whether they were born in the European Union ($\Phi^2 = 0.02$) and global region of birth ($\Phi^2 = 0.02$).

The Internet national respondents were also very similar to the Natsal national sample in terms of self-assessed health status ($\Phi^2 = 0.09$), HIV testing ($\Phi^2 = 0.01$), drinking alcohol ($\Phi^2 = 0.09$) and injecting drugs ($\Phi^2 = 0.01$). This is despite the differences in how the variables were measured, as described earlier. These national samples showed remarkable similarity in likelihood of testing for HIV (52.9% of Natsal, 52.8% of Internet 2003 and 53.4% of Internet 2002), frequency of drinking alcohol and whether they had injected drugs (99.1% of Natsal said "no" vs 98.2% of Internet 2003). Natsal and the Internet 2003 samples both reported that their first sex with another male, with or without genital sex, had been at 18 years old (s.d. 6.1 and 5.4, respectively).

The Internet national respondents were more likely to report having STDs in the last year (vs 4.8% of Natsal vs 16.9% of Internet 2003, $\Phi^2 = 0.32$; and 19.0% of Internet 2002, $\Phi^2 = 0.44$) and while the reporting of gonorrhoea, syphilis and "other" STDs was statistically similar ($\Phi^2 = 0.01$ to 0.13), Internet respondents were more likely to report all three (gonorrhoea: 5.4% vs 1.3%; syphilis: 1.3% vs 0.0%; other: 14.4% vs 4.2%).

The likelihood of having had anal sex increased among the Natsal national respondents from 51.8% to 69.3% when the period of interest increased from 4 weeks to 6 months. Although the Internet samples were still more likely to have had anal sex over the previous 3 months than Natsal respondents over the previous 6 months (76.9% of Internet 2003 and 76.0% of Internet 2002), these distributions were statistically similar ($\Phi^2 = 0.03$ and 0.02 respectively).

6.3.2 Comparison of samples of London men

Comparisons between the London men from Natsal 2000 and the two Internet London samples are shown in Tables 6.6.1 to 6.6.8 (on the following pages).

Although the Internet London samples were very similar to the Natsal London sample in age, there was an indication that Internet respondents were slightly younger than Natsal respondents (mean ages 31.9 (s.d. 6.7) for Natsal vs 31.2 (s.d. 7.0) for Internet 2003 and 31.6 (s.d. 6.6) for Internet 2002). When respondents were placed into age groups, the values of phi-square were close to zero for both comparisons ($\Phi^2 = 0.00$ and 0.01), again showing that the Internet London samples were of a similar age to the Natsal London sample.

All three London samples were made up of a majority of white men, with the Internet samples statistically similar to the Natsal sample in terms of ethnic composition ($\Phi^2 = 0.10$ and 0.11). The Internet samples were, however, somewhat less ethnically diverse than the Natsal sample (77.9% of Natsal were white vs 89.9% of Internet 2003 and 90.9% of Internet 2002) and had very low proportions of black respondents (8.7% of Natsal vs 1.2% of Internet 2003 and 1.5% of Internet 2002).

Table 6.6.1: Comparison of samples of London men

	Natsal	Internet 2003	Internet 2002
Mean age (yrs)	31.9	31.2	31.6
Standard deviation	6.7	7.0	6.6
Base (n)	*38*	*493*	*1112*
t statistic		0.6	0.3

Note: *p ≤ 0.05, **p ≤ 0.01

Age group (%)

18 to 19	0.0	4.1	3.0
20 to 24	18.3	17.6	14.5
25 to 34	45.6	42.0	46.0
35 to 44	36.2	36.3	36.6
Base (n)	*34*	*493*	*1112*
phi-square		0.01	0.00

Note: categories "18 to 19" and "20 to 24" are collapsed to calculate phi-square

Ethnicity (%)

White	77.9	89.9	90.9
Black African	0.8	0.0	
Black Caribbean	7.9	0.8	1.5
Other Black	0.0	0.4	
Asian	4.2	2.3	2.8
Mixed / Other	9.1	6.6	4.8
Base (n)	*38*	*487*	*1112*
phi-square		0.10	0.11

Note: black respondents are collapsed into one category to calculate phi-square

Table 6.6.2: Comparison of samples of London men

	Natsal	Internet 2003	Internet 2002
Work status (%)			
Working	71.9	80.1	80.7
Not working	17.3	9.1	12.1
Student	10.8	10.8	7.3
Base (n)	*38*	*493*	*1112*
phi-square		0.05	0.04
Highest educational qualification (%)			
Degree level	46.6	60.8	-
Not degree level	53.4	39.2	-
Base (n)	*38*	*492*	-
phi-square		0.08	-
Number of post-16 years in education (%)			
None	10.2	11.0	10.7
Up to 2 years	20.2	18.4	15.3
3 years or more	63.9	63.7	66.7
Still in f-t education	5.7	6.9	7.3
Base (n)	*38*	*490*	*1112*
phi-square		0.00	0.02

Table 6.6.3: Comparison of samples of London men

	Natsal	Internet 2003	Internet 2002
Social class (%)			
I	8.6	10.0	-
II	57.4	57.4	-
IIIN	18.2	21.8	-
IIIM	12.2	8.1	-
IV	1.9	2.7	-
V	1.7	0.0	-
Base (n)	*33*	*408*	-
phi-square		0.04	-
Born in the United Kingdom or Eire (%)			
Born UK/Eire	66.3	76.7	-
Not born UK/Eire	33.7	23.3	-
Base (n)	*38*	*486*	-
phi-square		0.05	-
Born in the European Union (%)			
Born EU	71.5	84.4	-
Not born EU	28.5	15.6	-
Base (n)	*38*	*486*	-
phi-square		0.08	-

Table 6.6.4: Comparison of samples of London men

	Natsal	Internet 2003	Internet 2002
Global region of birth (%)			
Europe – EU	71.5	84.4	-
Europe – non-EU	3.9	2.1	-
Australia, NZ	2.1	2.5	-
N America	0.0	1.9	-
S & Central America	5.3	1.2	-
Caribbean	6.1	0.6	-
Indian subcontinent	3.3	0.4	-
China	0.0	1.4	-
Other Asia	1.4	1.6	-
Middle East, N Africa	0.7	0.6	-
W Africa	2.5	0.0	-
Central and E Africa	0.0	1.9	-
South Africa	3.1	1.4	-
Base (n)	*38*	*486*	-
phi-square		0.08	-

Note: categories are collapsed into "Europe", "Australia, NZ", "America and Caribbean", "Asia" and "Middle East and Africa" to calculate phi-square

Self-assessed health status (%)			
Poor	2.4	4.1	-
Fair	3.7	15.6	-
Good	41.1	27.8	-
Very good	52.8	52.5	-
Base (n)	*34*	*493*	-
phi-square		0.44	-

Table 6.6.5: Comparison of samples of London men

	Natsal	Internet 2003	Internet 2002
Whether tested for HIV (%)			
Yes	52.9	66.7	68.0
No	47.1	33.3	32.0
Base (n)	*37*	*493*	*1112*
phi-square		0.08	0.09
Frequency of drinking alcohol in last year (%)			
1 or 2x a week	80.3	70.4	-
1 or 2x a month	10.2	19.4	-
1 or 2x a year	1.9	4.1	-
Never	7.6	6.1	-
Base (n)	*34*	*490*	-
phi-square		0.12	-
Whether injected drugs in last year (%)			
No	99.2	97.3	-
Yes	0.8	2.7	-
Base (n)	*38*	*489*	-
phi-square		0.05	-

Table 6.6.6: Comparison of samples of London men

	Natsal	Internet 2003	Internet 2002
Mean age of first sex with a male			
Age in years	17.4	17.1	-
Standard deviation	6.9	5.1	-
Base (n)	*38*	*493*	-
t statistic		0.2	-

Note: comparison of "sex of any kind" (Internet and HIV) with "any sexual experience or contact (Natsal); *p ≤ 0.05, **p ≤ 0.01

Mean age of first (genital) sex with a male			
Age in years	17.6	17.1	-
Standard deviation	6.9	5.1	-
Base (n)	*36*	*493*	-
t statistic		0.5	-

Note: comparison of "sex of any kind" (Internet and HIV) with "sex involving genital contact (Natsal); *p ≤ 0.05, **p ≤ 0.01

Whether had STDs in last year (%)			
Yes	7.9	21.5	25.1
No	92.1	78.5	74.9
Base (n)	*38*	*489*	*1112*
phi-square		0.25	0.41

Table 6.6.7: Comparison of samples of London men

	Natsal	Internet 2003	Internet 2002
Whether had gonorrhoea in last year (%)			
Yes	3.9	-	8.7
No	96.1	-	91.3
Base (n)	*38*	-	*1112*
phi-square		-	0.06
Whether had syphilis in last year (%)			
Yes	0.0	-	3.1
No	100.0	-	96.9
Base (n)	*38*	-	*1112*
phi-square		-	-
Whether had another STD in last year (%)			
Yes	6.0	-	17.6
No	94.0	-	82.4
Base (n)	*38*	-	*1112*
phi-square		-	0.24

Note: phi-square cannot be calculated because no cases were expected to have had syphilis in the last year

Table 6.6.8: Comparison of samples of London men

	Natsal	Internet 2003	Internet 2002
Whether has had recent anal sex (4 weeks vs 3 months) (%)			
	4 weeks	3 months	
Yes	41.8	81.9	82.4
No	58.2	18.1	17.6
Base (n)	*38*	*493*	*1112*
phi-square		0.66	0.68

Note: comparison of anal sex in the past 4 weeks (Natsal) with the past 3 months (Internet and HIV)

Whether has had recent anal sex (6 months vs 3 months) (%)			
	6 months	3 months	
Yes	58.5	81.9	82.4
No	41.5	18.1	17.6
Base (n)	*38*	*493*	*1112*
phi-square		0.23	0.24

Note: comparison of anal sex in the past 6 months (Natsal) with the past 3 months (Internet and HIV)

These London samples were similar in terms of work status (Φ^2 = 0.04 and 0.05) but the Natsal respondents were somewhat less likely to be working than Internet respondents (71.9% of Natsal vs 80.1% of Internet 2003, 80.7% of Internet 2002).

The samples were statistically well matched on both of the educational variables (Φ^2 = 0.00 to 0.08) but the Internet 2003 respondents were rather more likely to have been educated to at least degree level (60.8% vs 46.6% of Natsal). Although the Internet samples might have been expected to report more years of post-16 education because of the discrepancy in the questions measuring this, the Natsal and Internet samples were remarkably similar in terms of number of years in post-16 education.

The Internet 2003 London sample was also very similar to the Natsal London sample in terms of social class (Φ^2 = 0.04), with the majority of respondents in social classes I and II (66.0% of Natsal vs 67.4% of Internet 2003).

The Internet 2003 London respondents were rather more likely to be born in the UK or Eire than the Natsal London respondents (76.7% vs 66.3%) and to be born in the European Union (84.4% vs 71.5%), although the differences between the samples was not statistically strong (Φ^2 = 0.05 and 0.08 respectively). There were also small differences between respondents in the two samples on their global region of birth (Φ^2 = 0.08), with the Internet sample more likely to be born in the EU and the Natsal sample more likely to be born in south and central America (5.3% vs 1.2%), a Caribbean country (6.1% vs 0.6%), India, Pakistan or Bangladesh (3.3% vs 0.4%) and west Africa (2.5% vs 0.0%).

The Internet 2003 London and Natsal London respondents were not statistically well matched on self-assessed health status (Φ^2 = 0.44). Internet 2003 respondents were more likely to report poorer health, with 19.7% rating their health as poor or fair compared to 6.1% of Natsal respondents. This is unlikely to be explained by the inconsistent scales used to measure health status which would have predicted differences in the opposite direction.

Internet London respondents were rather more likely to have been tested for HIV (52.9% of Natsal vs 64.9% of Internet 2003 and 68.0% of Internet 2002), although the samples were statistically similar ($\Phi^2 = 0.08$ and 0.09).

These London samples were statistically similar in terms of alcohol consumption ($\Phi^2 = 0.12$) and drug injecting ($\Phi^2 = 0.05$), although Natsal respondents were somewhat more likely to report drinking alcohol at least once or twice a week (80.3% of Natsal vs 70.4% of Internet 2003). This is consistent with the different scales used which would predict higher levels of reported drinking in the Natsal sample. Injecting drugs was of low prevalence in both samples that measured this behaviour.

For both Natsal and the Internet 2003 samples, the age of first sex with another male was 17 years old (s.d. 6.9 and 5.1, respectively), although the age of first genital sex was closer to 18 years for Natsal respondents.

Respondents from both Internet London samples were considerably more likely to have had a sexually transmitted disease (STD) in the last year (7.9% of Natsal vs 21.5% of Internet 2003, $\Phi^2 = 0.25$; 25.1% of Internet 2002, $\Phi^2 = 0.41$). The Internet 2002 sample reported consistently higher proportions of gonorrhoea, syphilis and other STDs and the greatest difference was for "other" STDs (6.0% of Natsal vs 17.6% of Internet 2002, $\Phi^2 = 0.24$).

The likelihood of having had anal sex increased among the Natsal respondents from 41.8% to 58.5% when the period of interest was increased from 4 weeks to 6 months. However, Internet respondents were much more likely to have had anal sex over the previous 3 months than Natsal respondents over the previous 6 months (81.9% of Internet 2003, $\Phi^2 = 0.23$; 82.4% of Internet 2002, $\Phi^2 = 0.24$).

6.3.3 Comparison of samples of men from outside London

Comparisons between the Natsal 2000 and the Internet samples of men from the regions are shown in Appendix 6.

Internet respondents from outside London were significantly younger than Natsal respondents from outside London (mean age 31.9 (s.d. 6.0) for Natsal vs 28.8 (s.d. 7.7) for Internet 2003 and 29.4 (s.d. 7.4) for Internet 2002, $p \leq 0.01$). It is also notable that while the Natsal respondents from outside London had the same mean age as their counterparts from within London, the Internet respondents from outside London were slightly younger than their London counterparts. The difference in the ages of the Natsal and Internet samples from outside London was also reflected in the comparison of age groups ($\Phi^2 = 1.01$ and 0.65). The vast majority of men in all three samples were white (99.1% of Natsal vs 97.4% of Internet 2003, $\Phi^2 = 0.04$; and 96.6% of Internet 2002, $\Phi^2 = 0.07$), with notably less ethnic diversity than the London samples.

There were clear differences between the Internet samples from outside London and the Natsal sample in terms of work status ($\Phi^2 = 1.16$ and 0.42), with the former much more likely to be students (2.5% of Natsal vs 19.2% of Internet 2003 and 12.6% of Internet 2002) and the latter more likely to be working (85.5% of Natsal vs 74.8% of Internet 2003 and 77.4% of Internet 2002).

The Natsal and Internet 2003 respondents were similarly likely to be educated to degree level (42.2% vs 38.2%, $\Phi^2 = 0.01$), but both Internet samples were more likely to still be in full time education (2.8% for Natsal vs 14.5% for Internet 2003, $\Phi^2 = 0.55$; and 12.6% for Internet 2002, $\Phi^2 = 0.39$) and less likely to have had no post-16 years in education (27.3% for Natsal vs 15.7% for Internet 2003 and 18.2% for Internet 2002).

The Internet respondents were statistically similar to the Natsal respondents in terms of the type of area where they lived ($\Phi^2 = 0.00$ and 0.02) with the majority of respondents in all three samples lived in urban areas (90.7% of Natsal vs 89.7% of Internet 2003 and 86.3% of Internet 2002). The distribution of the Internet 2003 respondents around the country was also similar to that of the Natsal respondents ($\Phi^2 = 0.18$).

The Internet 2003 and Natsal samples were reasonably well matched in terms of social class ($\Phi^2 = 0.21$), although Natsal respondents were more likely to be in social classes I and II (67.2% vs 56.4%) and less likely to be in social class IIIN (11.0% vs 23.9%).

The Internet respondents from outside London were very well matched with the Natsal respondents in place of birth ($\Phi^2 = 0.00$ and 0.02), with over 90% of men in both samples born in the UK and Eire or in the European Union. No comparison was therefore made of global area of birth. The Internet respondents were also remarkably similar to the Natsal respondents in terms of self-assessed health status, HIV testing, drinking alcohol and injecting drugs, with values of phi-square close to zero on all counts ($\Phi^2 = 0.00$ to 0.04). Mean age of first sex with another male, with or without genital sex, was 18 years old for both Natsal samples and Internet 2003 (s.d. 5.7 and 5.4 respectively).

Internet respondents were much more likely to report having STDs in the last year (3.3% of Natsal vs 15.2% of Internet 2003, $\Phi^2 = 0.44$; and 16.6% of Internet 2002, $\Phi^2 = 0.55$). Although none of the Natsal respondents reported gonorrhoea or syphilis, 4.1% and 0.6% of Internet 2002 respondents did so. Internet 2002 respondents were also more likely to have had other STDs (3.3% of Natsal vs 13.2% of Internet 2002, $\Phi^2 = 0.31$).

When the period of interest was increased from 4 weeks to 6 months, the likelihood that Natsal respondents had had anal sex increased from 56.5% to 74.4%. Although the Internet samples from outside London were more likely to have had anal sex over the previous 3 months than Natsal respondents over the previous 4 weeks (75.2% of Internet 2003 and 73.6% of Internet 2002), the samples were statistically similar (0.14 and 0.12).

6.3.4 Comparison of national gaydar and gay.com samples

Comparisons between the 2003 national samples of men from gaydar and gay.com are shown in Appendix 7.

When the Internet 2003 national sample was divided into those who came to the survey from the gaydar website and those who came from gay.com, gaydar respondents were found to significantly younger than gay.com respondents (29.2 (s.d. 7.7) vs 30.7 (s.d. 7.3), $p \leq 0.01$). This was also reflected in the comparison of age groups ($p \leq 0.01$).

The gaydar and gay.com samples were significantly different on ethnic composition, with gaydar comprising fewer Asian men (0.5% vs 1.8%) and men of other or mixed ethnicity (2.8% vs 4.9%, $p \leq 0.05$). The samples were very similar on work status but gaydar respondents were less likely to be educated to degree level (44.3% vs 51.1%, $p \leq 0.05$). This difference was not reflected in the number of years of post-16 education.

The two samples were similar on whether they lived in an urban or rural area. Although they were fairly similar in terms of geographical distribution about the country, the samples were significantly different on this variable. The greatest difference was a higher proportion of gay.com men living in Scotland (14.1% vs 9.1%, $p \leq 0.05$). The samples were well matched on social class and there was no significant difference between the samples on global region of birth but gaydar respondents were more likely to be born in the UK or Eire (90.5% vs 84.8%, $p \leq 0.01$) and in the EU (93.2% vs 90.0%, $p \leq 0.05$).

The samples were similar in terms of self-assessed health status, HIV testing, drinking alcohol, injecting drugs and age of first sex with another male.

The Internet respondents were asked about sexual partners in the past year – whether they were all male, mostly male, male and female, mostly female or all female – whereas Natsal asked about the respondents' sexual partners during their lifetime. Because lifetime experience is very different from recent experience and may include transient past experiences (Mercer *et al.* 2004), this variable was not included in the above comparisons with Natsal. This comparison of the gaydar and gay.com samples, however, found that gay.com respondents were more likely than gaydar

respondents to be behaviourally bisexual. 88.5% of gaydar respondents had had sex with only men in the past year compared to 81.5% of gay.com respondents ($p \leq 0.01$). It is interesting to note that 17.4% of the national sample of Natsal respondents reported that they had had sex with only or mostly women in their lifetime compared to 3.3% of Internet 2003 respondents over the past year[4], which may indicate that Natsal respondents are a more behaviourally bisexual group of men.

Gaydar respondents were more likely than gay.com respondents to report STDs in the last year (18.7% vs 12.8%, $p \leq 0.05$) and to have had anal sex in the past 3 months (78.8% vs 72.8%, $p \leq 0.05$).

[4] Note that the small number of men in each sample who reported that they had only had sex with women are included in these comparisons because they reported that they had had sex with at least one man in the past year elsewhere in the questionnaire.

6.4 Discussion of findings

This section presents a discussion of the above findings and concludes with their implications to survey researchers.

The Internet permits access to geographically dispersed samples of minority populations. In this way, the Internet and HIV study was able to collect data from four thousand gay and bisexual men (under the age of 45) across the UK in 2002 and two thousand men in 2003. This compares to the 171 men (before weighting) for whom data were collected in Natsal's 2000 national probability sample. The findings from the analysis undertaken here show that the online volunteers to the Internet and HIV study were distributed across the country in a similar way to the Natsal sample and were equally likely to live in an urban area.

In line with comparisons of online and offline samples from other populations, online convenience samples of gay and bisexual men have been found to be younger than offline convenience samples (Elford *et al.* 2004b; Ross *et al.* 2000). Such findings will of course depend on the comparison group and one sample of gay and bisexual men that was recruited in bars was found to be younger than the Internet sample (Rhodes *et al.* 2002). The data here show that the Internet national samples were significantly younger than the Natsal sample. The picture within London was slightly different, with the Internet London samples of very similar age to the Natsal London sample.

In Great Britain, the highest proportion of ethnic minority people live in London, where 29% of residents are from an ethnic minority (White 2002). This is reflected in the finding that among the gay and bisexual men from outside London, the vast majority of men in both Natsal and Internet samples were white. Thus, 7.7% of the Natsal national sample were from the ethnic minorities compared to 4.4% of the men in the Internet 2003 national sample. Although studies that have investigated the ethnic composition of online and offline samples have indicated a general reluctance of ethnic minorities to participate in either survey mode (Bailey *et al.* 2000; McCabe *et al.* 2002), these findings should be set in the context of the 2001 census which

found that 7.9% of the UK population were from ethnic minorities (Office for National Statistics Census 2001, http://www.statistics.gov.uk). Within London, the Natsal sample also contained a higher proportion of ethnic minority men than either Internet sample, the main difference being that Natsal drew a larger proportion of London black Caribbean men.

This difference between London and the rest of the country was also found in respondent place of birth. Outside London, the Internet 2003 samples were similar to the Natsal sample, with the majority of respondents born in the United Kingdom or Eire. Within the capital, two thirds of Natsal respondents were born in the UK or Eire compared to three quarters of Internet 2003 respondents. The Natsal London sample was more likely to contain men born in south and central America, the Caribbean, India, Pakistan and Bangladesh, and west Africa whereas the Internet London sample was more likely to contain men who were European born.

The results of previous studies that have compared the education of online and offline samples have not been consistent, although Internet samples of gay and bisexual men were found to have lower levels of education than offline convenience samples (Elford *et al.* 2004b; Ross *et al.* 2000). The data here show that Internet and Natsal national samples were well matched on measures of education, but the younger Internet samples from outside London were significantly more likely to be students. One fifth of Internet 2003 respondents from outside London described themselves as students compared to less than 3% of Natsal respondents. The findings provide evidence that Internet surveys of gay and bisexual men conducted outside London are more likely to attract a higher proportion of student volunteers than would be captured by a probability sample.

The level of education of the Internet London samples examined here was also found to be very similar to the Natsal sample. The Internet 2003 London respondents were rather more likely to be educated to degree level but the number of years in post-16 education was closely matched. This education profile is reflected in the work status of these London respondents. Both Internet samples were more likely to be working

than the Natsal sample but the proportion of students was similar to the Natsal sample.

A number of studies have indicated that Internet samples are more likely to be wealthy than their offline equivalents. The findings of the comparison undertaken here show that the majority of respondents in the Internet 2003 and Natsal samples came from non-manual social classes (I, II and IIIN) in London or outside the capital. Although these classes constituted 80% of the Natsal national sample and 82% of the Internet 2003 national sample, they made up only 46% of the general national male population of working age in 2000[5] (Office for National Statistics Labour Force Survey, Spring 2000, http://www.statistics.gov.uk). This indicates that the social class of the Internet sample was representative of the sub-sample of men drawn from a national probability sample but raises the issue of whether Natsal provided a representative sample of MSM from the manual classes.

The Internet 2003 sample tended to report health status in a similar way to the Natsal sample, although Internet respondents were more likely to say that their health had been "fair" rather than "good". This is contrary to the direction predicted by the use of different scales and may be indicative of genuine differences in perception of health status. The findings for frequency of drinking alcohol were broadly similar, but Internet 2003 respondents reported somewhat less frequent drinking than their Natsal sample counterparts both in London and the regions. This is in the direction predicted by the different scales that were used. The Internet sample did not differ from the Natsal sample in reporting drug injecting behaviour.

Differences between samples from within and outside London emerge again in relation to HIV testing. Outside London, just over half of the Natsal sample had tested for HIV, as had a similar proportion of the Internet samples. Within London, the proportion of men from the Natsal sample who had tested for HIV was again around half, whereas two-thirds of the Internet respondents had been tested. These

[5] It should be noted that 10% of this sample was coded as "other" whereas this category was not used in the above social class distributions for the Natsal or Internet 2003 samples.

findings indicate that the use of Internet samples to estimate HIV testing prevalence among gay and bisexual men in the capital may lead to over-estimation whereas such samples may be a better reflection of prevalence in the rest of the country.

Around one quarter of Internet London samples reported having had an STD in the last year compared to 8% of the Natsal London sample. Although all three samples reported fewer STDs outside the capital, the higher prevalence of STD reporting by the Internet samples was also found in the regions and this general trend was repeated for gonorrhoea, syphilis and other STDs both within and outside London. This is despite the fact that respondents in the Natsal sample may have reported STDs over a longer time period due to the different questions upon which these measures were based.

There was no difference between the Internet and Natsal samples in age of first homosexual experience. There were significant differences, however, in recent sexual practices. Over eighty percent of the Internet London samples had had anal sex over the past three months, whereas only sixty percent of the Natsal London sample reported anal sex in the previous six months. Interestingly, whereas the Natsal sample from outside London reported more anal sex in the previous four weeks and the previous six months than their Natsal London counterparts, the Internet samples from outside London reported less anal sex in the past three months than their Internet London counterparts. This resulted in three-quarters of Natsal respondents reporting anal sex in the previous six months, which was equivalent to the amount of anal sex reported by the Internet samples in the previous three months.

The data show that differences between the Internet and Natsal samples on reporting indicators of HIV risk behaviour (HIV testing, STDs and having anal sex) vary according to whether the Internet sample was recruited inside or outside London. Although it may be possible to apply weights to the Internet samples to iron out these differences, further exploration is required to examine Sparrow and Curtice's (2004) concern that weighting data may prove to be unsatisfactory.

Although the findings show a remarkable similarity between the Internet and Natsal samples on a number of variables, it should be noted that the use of Internet samples to estimate the prevalence of high risk sexual behaviour among gay and bisexual men in London may lead to over-estimation whereas there may be less bias associated with estimations based on data collected outside London. The sexual profile of the Internet samples is consistent with previous findings that individuals who volunteer for sex surveys have greater sexual experience (Bogaert 1996; Strassberg & Lowe 1995). It is also the case that the men in these Internet samples were recruited through online gay community contacts and this is likely to lead to greater sexual activity, as was found in other offline comparisons of probability and convenience samples (Meyer & Colten 1999; Sandfort 1997).

Previous findings have shown that different websites for meeting same sex sexual partners attract different men (Bull *et al.* 2004). The profile of gaydar and gay.com men was also found to be different in certain important respects. The gaydar respondents were slightly younger, less educated, less ethnically diverse and more likely to be born in the UK or Eire. They were less likely to be behaviourally bisexual and more likely to report STDs in the last year and anal sex in the previous three months. These findings provide support for the assertion that the Internet is not a homogenous location and that self-selected Internet samples will vary in the same way that offline convenience samples vary according to where they were recruited.

One of the limitations of the present study is that respondents were not randomly assigned to mode and as a result it is not possible to identify the impact of survey mode on sample composition. Furthermore, the variables used in the analysis were based on questions using different wording or response categories which is likely to have influenced comparability of responses. Although the main picture that emerged was that the Internet samples were more likely to report high risk sexual behaviour, the findings also suggest that the location where these samples are recruited has an important impact on sample composition. This was found to be the case in terms of geographical location (within or outside London) and web location (gaydar or gay.com). The findings indicate that different segments of the gay and bisexual

population may be reached via different sampling strategies (Sandfort 1997) and, with this in mind, the Internet provides an effective new venue for recruiting samples of gay and bisexual men.

6.4.1 Implications for researchers

The Internet national samples were broadly similar to the Natsal national sample on a number of sociodemographic and behavioural variables, with the notable exception of those indicating HIV risk behaviour. Differences were more pronounced, however, when samples were restricted to London alone. The findings indicate that the prevalence of high risk behaviour among gay and bisexual men based on self-selected Internet samples is likely to be over-estimated, particularly in London. Although Internet samples have an important role to play in establishing baseline estimates and monitoring trends in HIV risk behaviour, the data that they generate should be seen in this context. However, the Internet's capacity to reach a national sample of gay and bisexual men that may otherwise be difficult to capture makes use of the web survey an extremely attractive option for collecting data on trends in HIV risk behaviour among gay and bisexual men.