Readings on Sexual Diaries

Articles and chapters referred to in

INFORMATION ABOUT SEXUAL DIARIES
BIBLIOGRAPHY
Research note

'Something sensational . . .* The sexual diary as a tool for mapping detailed sexual behaviour

Tony Coxon

Abstract

Reliable information about the detail and context of sexual behaviour is crucial for describing high-risk behaviours for the transmission of AIDS/HIV, and for estimating parameters of epidemiological models. It is possible to adapt the method of systematic diary-keeping to this end and develop a tool which is also close to the subject’s conceptions of the ‘chunks’ of sexual behaviour.

The structure of reports of sexual behaviour is viewed as akin to linguistic structure, with a basic unit of the sexual act built into a ‘session’. Such structure lends naturally to a coding system which allows natural language accounts to be represented symbolically, and parsed and analysed formally. The coded version is easily learnt and used, and is incomprehensible to outsiders.

Examples are given of its use to collect systematic data on homosexual and bisexual behaviour.

Introduction

We here describe the use of a systematically-kept diary of sexual behaviour as a method of data-collection well-adapted to producing reliable and detailed data, especially for epidemiological studies.

Our domain of interest is the sexual behaviour of homosexual and bisexual men and its changes under the impact of AIDS/HIV. Despite the pioneering work of Kinsey, who used the sexual life-history collected in a face-to-face interview as his primary mode of data collection, most systematic studies aimed at describing gay sexual behaviour have relied upon the pre-coded self-administered

* 'I never travel without my diary. One should always have something sensational to read on the train'. Oscar Wilde, The Importance of Being Earnest.
everyday discourse among (British) gays. We believe this to be a crucial element in the success of its use in other groups and cultures.

2 The structure of sexual action

The sexual diary is an egocentric document, at least in the sense that it is written purely from the vantage-point of the subject, and this is a critical characteristic (and perhaps, limitation) of our version. Entries in the diary refer to sexual activity in which the subject (and his partner/s) was directly involved. Sometimes a subject's egocentric account will be corroborated (or even contradicted) by his partner's account if one is fortunate enough to have both of them keeping diaries independently, and this is fascinating information that might in some cases be used to 'repair' another account. (How else, for example, are we to know that a given orgasm was faked and therefore not risk behaviour?) Similarly, the subject may be able to give a perfectly coherent account of what his partners in a three-some did, but we would normally prefer the pairwise accounts of those directly involved, and we cannot realistically expect that a subject in full sexual arousal is likely or capable of giving accurate dispassionate descriptions of other people's behaviour at this point.

Our account of sexual behaviour, then, is a highly structured one, and owes much of its force to a linguistic analogy, which helps in its exposition.

2.1 Session and act

The basic meaningful unit we take as the 'sexual session' [sentence] fixed in time and usually in space, where sexual activity took place and the participant/s did not change. The session consists of a specific context in which sexual acts take place:

(i) context, details the time and place as well as relevant information about other partners, if any:

Tuesday, May 6th, at 10.30 p.m., with P1 at his flat, with P1 elsewhere identified as:

my primary partner, 36, accountant, who I live with.

in which (ii) one or more sexual acts occurred, detailing 'who did what and with what and to whom', as the Limerick accurately expresses it.

I wanked him, then he sucked me, and after that I sucked him off till he came.

Any of the sexual acts may have

'accompaniments' (such as condoms, lubricant/s, recreational drugs), so that the actual example of (ii) might be:

I wanked him while he sniffed poppers [nitrites], then he sucked me whilst I wore a condom. . . .

A few words on each component.

The context specification is necessary to set the scene and dramatis personae, locating the behaviour, possibly as part of a complex sequence of sessions, and disambiguating partner reference, so that the participants and the roles they adopt will be clear. The context also includes any preparations, such as watching a pornographic video, or having been out to a pub or club.

2.1.1 The sexual act

The sexual act (or acts) form the core content of a session, and constitute the words of the sentence. Each sexual act can be thought of as a verbal root (specifying what is actually done), with a prefix (indicating the modality of the act) and a suffix (indicating if and how - orgasm occurred).

Currently, the basic behaviours are primarily homosexually oriented, but can readily be extended simply to bisexuality and heterosexual acts. They consist of a set of common behaviours (masturbation, fellatio, an intercourse, vaginal intercourse), and a longer list of less common behaviours (ano-brachial insertion, femoral intercourse . . .), all of which are readily encoded by the first letter of the most common 'street' name (masturbation by Wanking, etc). Note that these include behaviours which may well not be genital stricto sensu (such as corporal punishment using belts) or which are 'safer sex' variants of more common genital acts, such as inter-femoral as a substitute for anal intercourse. The set of sexual behaviours is of course not closed, since we are limited only by physical capability and human imaginativeness.

The modality of the sexual act refers to the binary, directional relational specification of the act performed: 'who [does what] to whom' (of ego and alter), and we currently use five variants:

(1) SELF: Ego performs the sexual act on himself

(binary, asymmetric):

(2) ACTIVE: Ego performs the sexual act on/to Alter

356
The sexual diary as a tool for mapping sexual behaviour

cases of 'accompaniments' it can be ambiguous who wore the condom - an important fact! Suffice it to say here that the Crisco version, being more detailed, is also less ambiguous. The structure together with associated codes form the basis of our SIGMA/Sexual Behaviour Coding Scheme (Sigma), which we use throughout Project SIGMA research instruments, and not simply in the Sexual Diary.

There is clearly a trading relation between the complexity of such a scheme and the willingness of subjects to use it; moreover one would be foolish on logical or practical grounds to attempt completeness. At certain junctures it simply becomes counter-productive to add to the complexity, especially if the events that motivate change in the scheme are infrequent or occur in only a small subgroup. Indeed, the vast proportion of gay sexual behaviour is encompassed in three acts and three modalities!

Nonetheless, the rarity of the 'surprising' of an event is not a sufficient criterion for excluding it, since some rare behaviours (such as 'fisting' [ano-brachial insertion]) are, or may turn out to be, highly implicated in HIV transmission, even if subjects are initially unaware of them. Instructions to subjects make this clear, and encourage him to use natural language where necessary to supplement or supplant the letter-code.

2.3.1 The several codes

After several years of experimentation (including once abandoning the abbreviated letter-code for subjects' use), we now distinguish three levels of coding for the sexual diary:

(i) the (natural language, but structured) report of the subject (which may include differing amounts of Subject Code)

(ii) the (full) Subject Code, which is defined in Appendix 1 below, and the subject in Notes For A Sexual Diary. It is rarely used in complete detail by any subject. The letter-code symbols are as close as possible to being mnemonics of the relevant word, so it is quite easy to 'read' after a little practice. It is actually (but to the Subject, invisibly) parasitic upon:

(iii) the fully-defined 'Chriscode' version, which is the machine-held data script in one-to-one correspondence with the Cardbox Subject Code record, but is more explicit and efficient for sorting and search purposes. This is the representation operated on by the retrieval and analysis programs.

An example of the different variants of these codes is given in Appendix One. This written version (transcribed and slightly
Tony Coxon

1970s frequently involved (unprotected) sessions involving the repeated sequence: PFI + PFXO. It is not difficult to see how the advent of the HTLV-3 virus in such a context was easily spread. It is also worth commenting that unless sequences of sexual acts are mapped, simple frequencies are useless and misleading.

3 Notes for a Sexual Diary provides the subject with an introduction to the Sexual diary, and (if desired) to the Coding. It is available from Project SIGMA at cost.

References


Acknowledgements

I am very grateful to Phil Looker, now a sophisticate in this area due to his tireless devotion to in-putting Sexual Diaries, for his valuable advice and help throughout the project. Particular thanks, too, are due to Chris Mitchell, the progenitor of Chriscode and programmer for the special-purpose software.

Appendix One: Syntax of Diary Codes (Subject code; Sept 1987)

(subsession) ::= (event) | (session) (conjugator)
(session) ::= (subsession) | (conjugator)

The sexual diary as a tool for mapping sexual behaviour

(subsession) ::= (event) | (session) (conjugator)
(event) ::= (mode) (act) (respondent’s orgasm) (other’s orgasm) (conjugator)
(mode) ::= S | M | A | P | H
(act) ::= F | W | S | VF | FI | MA | BR | TF | CP | FG | RI | WS | TT . . .
(respondent’s orgasm) ::= O I X IP
(other subject’s orgasm) ::= O I X | P
(modifier) ::= null | (associated object list) (conjugator)
(object) ::= C | L | P | D | DR | T . . .
(conjugator) ::= space | & | +

Diary

WEEK BEGINNING / 19
(day) (month) (year)

Remember: Each entry should include:
Time | Place | Number code of other persons (if any), from Partner list | What happened: (who did what to whom), viewed from YOUR position. Including Who (if anyone) came to orgasm and whether the orgasm was ON or IN the person / and finally: Accompaniments [Condoms, Lubricants, Dildoes, Poppers, Toys, Drugs].

Monday

Met P3 at a nightclub in S/W London, went back to a friends place. We took some acid, and some joints, then I gave him a blow job where he came in my mouth, also I was fisted where I came over myself. Also we got into some s/m scenes later that morning, he fucked me and came inside me. We also used some poppers (English).
In recent years, medical science has provided us with considerable insight into the nature and consequences of Human Immunodeficiency Virus (HIV) infection and AIDS. Medical scientists and epidemiologists have also attempted to predict or model the extent to which HIV infection may become more prevalent over the next few years. In order to make these predictions, mathematical models are frequently constructed out of the factors or parameters which are believed to influence the development of a particular epidemic. Of particular importance within the modelling of HIV infection are parameters which identify both the number of sexual partners an individual has and particular sexual acts (primarily anal intercourse) as critical factors affecting the rate of transmission of HIV.

But social science too is needed in order to understand the role played by particular patterns of sexual behaviour in the epidemiology of HIV infection. There are good reasons to suppose that epidemiological analyses which take at face value the self-reports which individuals give concerning the frequency and types of sexual acts they participate in may be seriously flawed in the predictions they make. The social context in which data is collected as well as expectancies about the uses to which it is to be put both influence the reliability and validity of the self-reports that people give.

In this chapter, I will try to identify some of the methodological, technical and substantive factors which interact in the collection of data relating to male gay behaviour. An attempt will also be made to explore the implications of these issues for efforts to predict the rate at which HIV infection may spread. In particular, attention will be focussed on two crucial parameters in epidemiological studies of AIDS — the number of sexual partners and the prevalence of receptive anal intercourse. In exploring the issues associated with making reliable and valid estimates of these forms of sexual behaviour, reference will be made to research investigating changes in sexual behaviour of gay men within the context of AIDS. This project, Project Sigma — A longitudinal study of the sexual behaviour of homosexual men under the impact of AIDS — is concerned to investigate the current sexual lifestyles of gay and bisexual/committed men. By interviewing respondents in this study at six month intervals and by asking them to keep detailed sexual diaries in between these interview points, it is possible to detect and evaluate changes in their sexual and related behaviour.

Medicine and Social Science

Many issues to do with AIDS, HIV infection, sero-positivity and infectivity are by their very nature medically defined. But being HIV antibody positive or being diagnosed as having AIDS is more than simply a medical matter, since these conditions profoundly affect people, their relationships and lives. The consequences of HIV infection can therefore be as serious socially and economically as they can be in medical terms. The competences required to study these latter aspects are not medical. Indeed an analysis of the psychological, sociological and economic aspects of HIV infection and AIDS may even be hindered by medical attitudes and orientations since here the issues are of a different order to those usually encountered in medical practice.

But this is not simply a plea merely for a fuller consideration of the social consequences of HIV infection and AIDS. In response to the claim made by many doctors that ‘laypeople’ are unable to contribute directly to purely medical matters, it may be important to point out that, equally, medical scientists and clinicians rarely possess the competence necessary to fully understand social aspects of their human data — competences which stem primarily from training within the social sciences.

Three interdisciplinary and professional rivalries would not be so serious if they did not have consequences for determining priorities in the funding of AIDS-related research. The most obvious example of this can be found in health education and AIDS. It is ironic that whilst enormous effort and cost is given to ensuring that drugs and vaccines are monitored and evaluated, it is assumed that health education to change sexual behaviour needs no corresponding investment. Only after considerable pressure have self-help groups such as the Terrence Higgins Trust and Body Positive received any government funding, and then the amount they have received has been small in comparison to their needs and activities. People
have become accustomed to government action which is purely responsive 
to crises and which enthrones financial self-help as a prime virtue. But the 
irony is that even on these terms many aspects of present policy are 
counter-productive. Lack of funding now not simply means more deaths, 
but also unnecessary suffering and grossly more expensive healthcare costs 
in the future.

However, brave claims about the value of social science analysis with 
respect to AIDS are not enough. We need to demonstrate how medical 
issues can be illuminated by social science and how an ignorance of social 
science may impede or even nullify medical findings. In those aspects of 
medical science in particular which aim to develop theory from the 
self-reports of people about their own behaviour (and this includes 
activities as diverse as taking clinical life-histories and obtaining 
information on prevalence of certain behaviours, in addition to more 
obvious areas like medical counselling), approaches to data collection and 
analysis which ignore sociologically relevant insights are simply deficient; 
these areas need social science input, if they are to be scientifically credible.

HIV Infection, Epidemiology and AIDS

Nowhere is this need more obvious and important than in epidemiological 
studies, dependent as they are on behavioural observation, report and 
information. In the case of the epidemiology of AIDS in particular, we are 
dealing with a domain that is only slightly less delicate to enquire into 
than income. For much of the relevant data will of necessity relate to 
groups that are socially stigmatized (intravenous drug users, prostitutes, 
gay and bisexual men) and to sexual activities that are personal, sometimes 
illegal and often socially invisible. Taken together these issues are likely to 
make us question the reliability and validity of conventional data 
collection procedures, and manifestly require the use of sophisticated 
research skills.

Instead of this, in most epidemiological studies of HIV infection, 
sexual behaviour and AIDS we typically find a heavy reliance on data 
collected in clinical settings, combined with an unrealistic willingness to 
extrapolate from dubious American studies of sexual behaviour conducted 
within from the pre-AIDS era. In this chapter I do not intend to discuss in 
any detail the distortions which can be introduced into our understandings 
of sexual behaviour by inappropriate sampling techniques nor by the social 
dynamics of the interview situation. Nevertheless, we need to recognize 
that sexual behavioural studies of those who already have a history of 
sexually transmitted disease (a persistent and clearly powerful co-factor in 
the etiology of AIDS) are unlikely to provide an unbiased picture of the 
behaviour of the entire at-risk population. Furthermore, being interviewed 
within the medical context of a clinic, by an interviewer who may be far 
from informed and possibly homophobic, augurs ill for truth-telling.

Reliability and Self-report Data

More disturbing in many ways is the clear and repeated evidence from 
research which suggests that in general the reliability of data gathered in 
such studies of sexual behaviour may be low (Coxon, 1986a; McManus and 
McEvoy, 1985). This is especially true of the very detailed information 
necessary for sophisticated epidemiological modelling. The main points 
concerning this lack of reliability can be summarized schematically as 
follows.

First, we rely on people’s willingness to report sexual behaviour to 
infer information about its quality and quantity. This raises serious 
problems concerning the reliability of the estimates that people may give. 
Very few studies have attempted to triangulate upon the ‘same’ sexual 
events in order to assess the reliability of the claims that one of the parties 
involved may make. In our own work exploring the dynamics of gay 
behaviour we have used diary methods to examine the sexual practices of 
our respondents. On occasions where both sexual partners have been 
project subjects (either as a couple or contingently) and have completed 
daily sexual diaries, it has been possible to study triangulated accounts of 
these ‘same’ events. Preliminary findings concerning the reliability of data 
collected by this method suggests that the convergence between accounts is 
usually surprisingly good so long as the events concerned are recorded soon 
after their occurrence and the people reporting these are well-motivated 
and accept the guarantees of confidentiality and anonymity they have been 
given.

Second, studies of the reliability of self-reports of sexual behaviour 
show that in general, the retrospective recall of information tends to be 
selective, ordinarily distorted and unreliable. Research which attempts to 
identify when respondents first gained particular sexual experiences is 
likely to show the same biases. However, our research suggests that when it 
comes to investigating more recent events, it is possible to increase the 
reliability of the data by repeatedly asking the same questions or by 
implying them at several points in the course of the same interview. In 
linear regression analyses of respondents’ estimates of their reported daily
behaviour, however, we find systematic individual differences both in overall accuracy and in distortion. Knowing this, individual accounts can be compared; without it, individual comparisons can be highly misleading.

Third, in trying to track the detail and sequencing of sexual behaviour, time-lapse is also an important factor to take into account in assessing the reliability (in the test-retest sense) of the data. The further back in time an event occurred, the less reliable people’s reports become. In general, our research seems to suggest that people seem generally incapable of recalling accurately what happened a fortnight ago, let alone before this.

On these three sets of grounds we should be critical of the reliability of data emanating from routine inventories of sexual behaviour collected in a busy clinic for the treatment of sexually transmitted diseases. When this same data is subsequently used to provide parameter estimates of the frequency and nature of particular sexual acts for use in modelling the spread of infection within a particular population, there are grounds for further concern.

Reliably Estimating the Number of Sexual Partners

Nowhere are the consequences of low-quality data-collection more important than in assessing the reliability (and hence the validity) of the responses given to apparently innocuous questions about the ‘number of partners’ an individual may have had. On a priori grounds it might be imagined that those who are sexually exclusive/monogamous would have no difficulty answering such a question, those who have multiple (regular) partners might have some, and that the rest (the promiscuous, the sluts or the sexual athletes according to one’s point of view) might have a great deal. In fact two quite different issues are intertwined here: accuracy of recall on the one hand and definitional problems relating to what is meant by the term ‘sexual partner’ on the other.

Problems to do with accuracy of recall are easier to deal with, since they relate to the issues of reliability earlier discussed. In recalling sexual encounters, there certainly seems to be a recency effect (people remember best what has happened in the immediate past), but from our research the accuracy of their recall also depends on how varied a person’s sex-life is and how large or significant a fraction of it is comprised of one-off contacts. As in the case of reliability of reported sexual acts, we have been able in our work to check the estimates our respondents give of the number of sexual contacts they have against the actual number they record in their sexual diaries. When we do this, however, the news is not good since estimates of the number of partners (and hence of the rate of change in partners) seem to be least dependable when people make reference to their non-regular partners. Since their sexual behaviour is also least predictable for such contacts, and their partner’s antibody status is less likely to be known (and as there is a higher chance of alcohol and/or drugs being involved), there are important consequences here for attempts to estimate epidemiological parameters from such data. This is to put on one side entirely the issue of health risks.

But this is not all. If we look more closely at some of the questionnaires that have been to gain estimates of the number of sexual partners that gay men attending clinics for the treatment of sexually transmitted diseases may have had, we find that in many cases answers to questions about the ‘number of partners’ could not be given in an openended way. Rather, subjects are often provided with fixed pre-supplied categories. Significantly, these categories are rarely of equal-intervals, but follow an implicit power relation: small intervals at the least frequent end of the scale, working up to large ones at the most frequent end. They thus produce an artefact the evidence of ‘promiscuity’ so often uncritically quoted in the medical and other press (figure 1).

There are good grounds therefore to be suspicious on technical grounds of the reliability and validity of data collected in clinic situations concerning the number of sexual partners an individual has had. But there are other important issues here to do with cognitive and socio-semantic matters.

---

**Figure 1: Questions asked during a recent survey of male homosexual behaviour**

18. How many different male sexual partners have you had in the past year?

- None
- Less than 5
- 6-50
- 51-100
- 101-500
- 501-1,000
- 1,001 or more

19. How many different male sexual partners have you had in your lifetime?

- None
- Less than 5
- 6-30
- 31-100
- 101-500
- 501-1,000
- 1,001-2,500
- 2,501-5,000
- 5,001-10,000
- 10,001 or more
Even were we to provide potential respondents with an open-ended or evenly graduated scale of categories, it is important to know how much confidence the respondent gives to the estimate. In this respect, it is surprising how often crucial prefatory comments which indicate that respondents are unsure about the estimate they are about to make can be ignored in recording and analyzing the data. Bell and Weinberg (1978) are virtually unique in asking their interviewers to find out how respondents arrived at their estimated numbers of sexual partners by providing them with the following coded alternatives: ‘Rough guess/grossing up/reasonably precise number/exact number’.

Again, not surprisingly, we find that the reliability of these estimates is negatively associated with stated number of partners. There are exceptions: some people appear to carry around a mental diary which they can quote from ad libitum, and others clearly set themselves targets. But however celebrated the latter category of person may be in the world of literature, they tend to be rare in practice. In our research, we have been struck with how often a strategy of ‘grossing up’ is used for estimating anything over a week. As we have found for reported sexual behaviour, the week is usually the critical period of time which can be recalled quite well. When asked to provide estimates over a longer period, subjects will typically ‘multiply up’ their weekly figure. To the extent that this is true, what one gets in response to such questions is not a report on a longer period, but an extrapolation from one (perhaps atypical) week. It may well be that the responses people also give to questions about changes in their sexual behaviour over the last n weeks or months have exactly the same shortcomings.

But the most critical issue is: what do people actually mean when they talk of a ‘sexual partner’? (and, for that matter, what do researchers?). Without this knowledge we have no basis to make inter-individual comparisons. In particular, if a respondent’s conception of a ‘sexual partner’ does not conform to the researcher’s idea of what is meant by this term, then data collection and subsequent analysis will at best be biased and at worst meaningless.

In our own longitudinal studies of the sexual behaviour of gay men, we initially adopted as a tentative working definition of a ‘sexual partner’ — ‘anyone with whom you had any form of sexual experience except walking when no-one came’ — only to find that this failed to do justice to the range of definitions with which our respondents operated. Subsequent work has shown the variety of respondents’ definitions to be truly immense. At one extreme may be the man who enjoys sadomasochism (S & M) and who counts as any ‘sexual partner’ any other man who manifests an erection in an S & M scene. At the other may be men who only count as ‘sexual partners’ those with whom they have had anal intercourse to the point of orgasm. Between these two extremes lie many other variations. In the light of this, in studies attempting to gain reliable estimates of sexual behaviour, respondents and researchers need to explore carefully and systematically what each other mean when they talk of ‘sexual partners’ and ‘sexual acts’.

Modelling the Transmission of HIV Infection

How are these points relevant to the broader medical and sociological referred to earlier in this chapter? Given what we currently know about the processes involved in HIV transmission, epidemiological modelling of the likely spread of infection is crucially dependent on the availability of reliable and valid estimates of the prevalence of those acts most intimately connected with viral transmission. Two such parameters are the number of sexual partners an individual has and the extent to which receptive anal intercourse is practised. For reasons already discussed, reliable and valid estimates of these two parameters will be difficult, if not impossible, to come by using conventional data collection procedures.

But there are other problems too. Modelling the spread of an epidemic such as HIV infection (amongst gay men at least) may be further complicated by the fact that there is an asymmetry of partner exchange rates amongst gay men. Whereas within the heterosexual context the rates of partner exchange are equal between men and women, in the case of homosexual populations this is not so because gay men engage in anally insertive and anally receptive acts with different and unequal rates of partner exchange for each of these acts. It would be simpler from the point of view of epidemiological modelling if those engaging in anal intercourse would adopt a consistent role — and indeed for some time epidemiological models assumed that they did. The picture is further complicated by the fact that these two forms of sexual behaviour are associated with considerably different risks of infection. Knox (1986) for example has recently claimed that, ‘about 95 per cent of all known AIDS-infected homosexual and bisexual men are reported as A(nally) R(ective)’. Whilst statements like these are important in identifying the sexual acts which are particularly dangerous insofar as the transmission of HIV is concerned, they effectively presuppose that gay and bisexual men adopt a consistent role in anal intercourse: a claim which needs further consideration.
Kinds of Homosexual or Types of Sexual Behaviour?

Effective epidemiological modelling of HIV infection crucially depends on information about sexual behaviour obtained from non-clinical populations. Until very recently this information has been derived from what elsewhere I have identified as ‘a willingness to conjure information out of thin air or by extrapolation from dubious American studies of the pre-AIDS era’ (Coxon, 1986b). In this context it may be salutary to think about an incident some months ago when within the space of the same week two eminent medical experts proclaimed, and with no qualification, that ‘the’ fraction of homosexuals who were ‘passive’ was 90 per cent in one case and 50 per cent in the other. It can firmly be said that there are no reliable estimates of this sort currently available for Britain, not least because to obtain them we would have to pre-suppose that amongst gay men, behaviour is thus segregated. Certainly preliminary data from our own research suggests that a far higher fraction of men engage in both of these practices (but with different rates of partner exchange for each of them) than restrict themselves to one.

In obtaining estimates of key parameters within their epidemiological models, Knox and others rely primarily on studies by Kinsey (1948) and on Gagnon and Simon (1974) for their data. There are problems with this data however since not only was it collected some time ago but within a cultural context very different to that of modern Britain. Unfortunately, the estimates given for the prevalence of certain forms of sexual behaviour in these studies have a tendency to become accepted ‘truths’ or ‘facts’. As an example of this, we can consider Knox’s (1986) recent claim (taken accurately, it should be stressed from Gagnon and Simon’s work) that, ‘the AP (“active”) homosexuals are many times (for example × 20) as promiscuous as the AR (“passive”) homosexuals, and therefore much less (for example, × 0.05) frequent in this sub-population (of homosexuals)’ (p. 167).

Now the empirical basis for Gagnon and Simon’s original claim is tenuous to say the least, being based on estimates obtained from non-representative samples even within the American context. But here we can see what was originally a provisional estimate becoming part of received wisdom — being written into an epidemiological model as a key parameter predicting future patterns of HIV infection in Britain. Given our general lack of knowledge concerning the prevalence of particular types of sexual practices amongst members of non-clinical populations in Britain it is hard to present to counter such claims empirically but data from our own research on Project Sigma may in time allow this.

But where do such dangerously misleading images and understandings of male homosexual behaviour come from? Their recent origins can be traced back to psychiatric explanations which seek to understand this behaviour in terms of ‘conventional’ heterosexual practices. At their crux, such theories argued that all gay men simply identified with the female gender. Subsequent accounts identified two exclusive types: the socially visible ‘effeminate passive’ homosexual and the (unseen) ‘dominant active’ homosexual.

Following Biebert’s et al (1962) work in the early 1960s, these two types of homosexual men came to be equated with two types of sexual behaviour. It was they who coined the terms ‘insertor’ and ‘insertee’ to reflect this supposedly fundamental difference in sexual activity. Significantly of course the data which allow such claims to be made derive from clinical samples. It is important to recognize though that views like these were challenged by more sensitive observers at the time they were made. Westwood (1960), for example, was already insisting that all homosexual men could not be typified as exclusively active or passive and pointed out that a significant number of them regularly interchanged roles. These views were subsequently echoed in Harry and DeVall’s (1978) work.

Elsewhere I have argued that we should abandon the notion that gay men are predominantly anything, and adopt instead the more realistic assumption that gay males engage in a range of both types of behaviour (Coxon, 1986a and 1986b). This means abandoning the insertor/ee and active/passive typing of people, to reserve this distinguishing only for acts. Which of these acts individuals engage in depends critically upon the situation and the type of relationship they are in. In order to develop more sophisticated epidemiological models relating to HIV infection we need to know precisely what differentiates sexual behaviour in these various contexts.

Interaction Between the Number of Partners and Types of Sexual Acts

But that is not the end of the matter. As was pointed out earlier, epidemiological models quite rightly identify as key parameters affecting HIV transmission, and intercourse and rates of change of partners. But until recently many have failed to explore the crucial role of interaction between these two variables in determining the relative ‘safety’ of particular sexual acts. Recent research by Van Griensven et al (1986) in the Netherlands has begun to investigate some of the behavioural factors
discriminating between those who do and those who do not acquire HIV infection following particular patterns of sexual activity. The results are striking. Not just for anal intercourse, but in general, it is the interaction of the particular sex act and the number of partners (NP) that does the discriminating between whether or not a person becomes HIV infected, rather than either factor separately (figure 2).

Shown as an operating characteristic, it is clear that 'number of partners' accelerates in different ways depending on the risk of the behaviour concerned. Sometimes, indeed usually, it does this linearly with a varying slope. However, in the case of anally receptive (AR) acts it does this as a strong power function. The difference between AR x NP and AR x NP here is quite striking.

For those involved in health education there may be a number of lessons to learn from this data. Perhaps future health education messages should not simply urge gay men to use condoms when having anal sex, but should say quite plainly, 'If you are f*cked, you are at by far the greatest risk, and this risk increases dramatically with every different partner'.

Conclusions

In this chapter I have tried to do two things. First, efforts have been made to identify some of the social factors influencing the reliability and validity of data relating to gay male practices. I have been critical of the inferences clinicians and others have drawn from research based on clinical samples of gay men, and have argued that further enquiry is needed to identify the nature and variability of gay male practice. Data from Project Sigma — A longitudinal study of the sexual behaviour of homosexual men under the impact of AIDS — may in time provide this.

Second, I have tried to identify some of the problems inherent in the epidemiological modelling of HIV infection. Until we have reliable and valid estimates of various kinds of sexual behaviour it would seem likely that we shall continue to have difficulty in making predictions about the possible spread of infection and in devising health education programmes based upon sound knowledge about the relative 'safety' of particular sexual practices.

References


COXON, A. P. M. (1986a) Report on Pilot Study: Project on Sexual Lifestyles of Non-heterosexual Males, Cardiff, SRU.

COXON, A. P. M. (1986b) 'It all depends... the impact of context on patterns of homosexual behaviour', paper given at Medical Sociology Conference, University of York.


Tony Coxon


AIDS — A Trade Union Issue

Hugh Robertson

In this chapter I intend to explore why HIV infection and AIDS is an issue for the British trade union movement by identifying some of the ways in which different trade unions have responded to AIDS so far. I will focus in particular on the rather different factors which have led two major trade unions to make very different responses to AIDS and AIDS-related issues. Finally, I hope to show that the trade union movement could play a greater role in tackling the issues raised by HIV infection and AIDS than it has done so far.

Initially, however, it is important to put discussion of these issues in context. Trade unions are far from peripheral to any debate concerning the social aspects of AIDS. The British trade union movement plays a unique role in British society. Proportionately it is the largest and strongest trade union movement in the non-communist world. Despite assaults on trade union rights over the last seven years, mass unemployment and the defeat of the miners’ strike, the trade unions are still a major social force, with in excess of ten million members and negotiating rights in most major industries and services. They have also, to some extent, been responsible for much of the progressive social change that has occurred over the last century.

What therefore are the implications of AIDS for this movement? In this chapter four separate, but interconnected, areas of trade union interest will be discussed. These concern conditions of service, health and safety at work, equal opportunities, and National Health Service resource provision.
SEXUAL BEHAVIOUR AND RISKS OF HIV INFECTION
Proceedings of an international workshop supported by the European Communities

EDITED BY
Michel HUBERT

PATTERNS IN HOMOSEXUAL RELATIONS: THE USE OF THE DIARY METHOD

P. DAVIES & T. FOXON

Bruxelles
Publications des Facultés universitaires Saint-Louis
Boulevard du Jardin botanique 43
1990
Patterns in Homosexual Relations: the Use of the Diary Method

by

Peter DAVID** and Tony COXON**

The HIV epidemic forces all of us who are sexually active to consider and perhaps to change our sexual behaviour. These changes are potentially of the most profound kind. Much health education that urges these changes has been couched in terms of simple prescriptions or proscriptions: don't fuck; always use a condom. More recently, and with more accurate knowledge of the modalities of transmission of HIV, attempts have been made to introduce generative messages, messages that contain rules that give the individual the means by which to decide whether a particular act is unsafe. Messages such as 'on him not in him' or 'don't exchange body fluids' are generative in this sense. As sexual behaviour changes in the face of the spread of HIV and the novelty of raspberry flavoured condoms wears off, the need for individuals and groups radically to re-evaluate the ways in which they have sex will become increasingly vital.

Elsewhere (Davies, 1989), I have drawn the analogy with changing diet. If the unsafe act is not particularly important to the individual's sexual appetite like, say, jam of truffles in the diet, then life without them is not a particular hardship, but if the acts are central to the process of sex, as say meat is to the non-vegetarian diet, then the change will be a more difficult one. As anyone who has changed from a meat-eating to a vegetarian diet will know, the change involves a complete re-appraisal of the process of eating, involving a move from heuristic to explicit rules for meal construction. Similarly, long-term changes of sexual practice to encompass safer sex will involve the greater or lesser re-appraisal of sexual activity, sexuality and sexual identity.

Since we now know that the unsafe acts are vaginal and anal intercourse, we must ask as scientists and as sexual individuals what is the place of these acts in the sexual "meal". If it is the optional extra, like the Parmesan cheese or the after dinner mint, then simply to omit it from the repertoire will not be difficult and simple prescriptive messages may suffice, but if it functions more like the filet mignon or the lobster thermidor, that is as the piece de resistance of the meal, then the omission will occasion more pain and heart searching and more complex messages and strategies will be required.

Responses to the risk of HIV infection may be classified into two broad types. The first, which involves an individual decision, is a decision to avoid sexual contact: to become celibate or monogamous or, more subtly to avoid those situations in which experience suggests that unsafe sex occurs. The second strategy, which, by contrast, involves negotiation, is an agreement not to engage in high-risk activities or to use a condom or whatever. It is important to note that these two types involve essentially different processes of response. The first is an individual decision and will be influenced by peer group pressure, cultural and sub-cultural norms and individual reserves of will-power but the second, particularly when with occasional partners, involves an actor in a negotiation into which obtrude issues of personal efficacy, power and responsibility in situations which themselves implicitly or explicitly channel power in particular ways.

Two linked questions arise. Firstly are there particular physical, symbolic or situational contexts in which unsafe sex is more likely to occur than in others? Secondly, are there identifiable types of sexual interaction which are more likely to include unsafe acts than others? This paper is an attempt to begin an answer to the second of these two questions.

In order to gain some leverage on this question, Project SIGMA has developed the method of the sexual diary (see Coxon, 1988 for a summary). The method gives a uniquely detailed description of the sequences of sexual action and allows the analysis of the influence of context (both physical and situational) of behaviour.

I. The sexual diary

It is important to recognise that the logical primitives for the study of sexual behaviour in general and the transmission of HIV in particular are sexual acts, not sexual preference, partners or sexual identity. HIV has a differential probability of transmission for each act, which is why we must be careful to refer to high risk activities rather than high risk groups or

* Project SIGMA, Department of Social Sciences, South Bank Polytechnic, London.
** Research Centre on Microsocial Change, University of Essex, Colchester.

1. Project SIGMA (Socio-Sexual Investigations of Gay Men and AIDS) is a linked set of research projects, funded by the Medical Research Council and the Department of Health whose assistance is gratefully recognised.
individuals. The diary method relies on a coding of sexual acts, both for the case of the diarist and for succinct computer representation. The appendix contains details of the code as presented to the prospective diarist. This is slightly different from and slightly less complex than the version of the code used for the computer record and diarists are encouraged, in the first instance, to keep their diaries in written or spoken English.

The acts are identified using a one- or two-letter mnemonic, drawn from the vernacular or street English terms for them. (Incidentally, we found that almost without exception, our respondents preferred to use these street terms rather than the more scientific Latin terms.) Table 1 gives a list of the codes together with their various translations.

Table 1: A Partial List of Sexual Activity Codes

<table>
<thead>
<tr>
<th>Mnemonic</th>
<th>Street</th>
<th>Scientific</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>Wanking</td>
<td>Masturbation</td>
</tr>
<tr>
<td>F</td>
<td>Fucking</td>
<td>Anal Insertion</td>
</tr>
<tr>
<td>R</td>
<td>Rimming</td>
<td>Oro-anal Contact</td>
</tr>
<tr>
<td>D</td>
<td>Douching</td>
<td>Douching</td>
</tr>
<tr>
<td>S</td>
<td>Sucking</td>
<td>Fellatio</td>
</tr>
<tr>
<td>VF</td>
<td>Fucking</td>
<td>Vaginal Insertion</td>
</tr>
<tr>
<td>FI</td>
<td>Fingering</td>
<td>Digital insertion</td>
</tr>
<tr>
<td>Etc.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These acts are modulated by prefixing to the code letter/s a single letter modifier drawn from the set S, P, A, standing respectively for Self, Passive and Active. Thus each act may occur, logically if not physically, in one of the three modalities: each may be done to myself, to me by another, or by me to another.

A further modifier M, standing for mutual, is also sometimes useful as a shorthand for simultaneous A and P. It is important to note that the modality modifiers follow the grammatical sense of the English terms that are indicated by the codes. Thus, the active-passive distinction in the codes does not follow the insertor-insertee distinction (compare for example AF and AS).

Where orgasm may occur as the direct result of one of the acts, a further two letters are added to the code, the first indicating the destination of the diarist's ejaculation, the second the destination of the partner's. The codes in each case are one of the letters O, I, C, E, X, which stand, respectively for On a person, In a person, into a Condom, Elsewhere or that there was no ejaculation.

The next and most important level of description of sexual activity is the session. A session is simply a succession of sexual acts taking place in sequence with the same partner or partners. Acts that occur simultaneously are joined with an ampersand. While in most cases the end of a session is fairly easy to distinguish, there is some residual difficulty in deciding how long a pause in a session has to be before it becomes a period of time between two sessions.

II. Some problems

Clearly, the diary method, while providing detailed and extensive information about sexual activity, carries along with it some problems.

A. Chunking

It is clearly unrealistic and undesirable that every physical movement be encoded in the diary. Rather, what happens is that people record activity in "natural" chunks. This suggests that there exists a language for talking about sexual activity which the diary method can capture. While this language may not be logically water-tight, this is no different from any other language. As Wittgenstein among many others has commented, language is fuzzy at the edges.

The problem with chunking is that slight differences in sessions may be lost and conflated into stereotypical sequences if the period between the action and the recording of it is at all extended. Although this is a major difficulty if we expect the method to record the reality of the activity, this is not the case if we are interested in the way that people construe sexual sessions. In seeking to influence or to change behaviour, it is the expectations that people have about sexual activity that are almost as important as the actuality.

B. Recall

As was suggested earlier, although the diary method alleviates the problem of recall in the study of sexual behaviour, it does not solve it. Although diarists are asked to make a record of their sexual activity as soon after the fact as is practicable, it is recognised that the exigencies of the situation may make the prompt recording of details impossible. We surmise from all the work done on recollection and memory that the longer the delay between act and record, the grosser the chunks that will be recorded and the more likely it is that the activity will be conflated to a stereotypical pattern.
C. Scenes and scenarios

The method as currently construed is adequate for the description of the physical acts of sexual behaviour. It is less well-adapted to the description of the settings within which that activity takes place. If we take seriously the notion of a language of sexual activity, then we must also consider the indexicality of that language, that is to say the likelihood that the meaning of particular acts or sequences of acts will differ according to the circumstances in which they take place. These circumstances may be physical or psychological. That is to say that particular acts may be more appropriate to certain places (compare, for example, the bed and the outdoor cruising ground) and also to the explicit or implicit scenario of the action. It is common, for example for SM sexual activity to make use of explicitly negotiated scenarios within which sexual activity occurs. Other sexual activity may also take place within less well-defined, less explicit, unnegotiated (that is to say taken for granted) scenarios.

D. Literacy

Although provision is made in the notes that explain the use if the diary (as in the appendix) for diarists to send in verbal records of their activity on cassette tapes, the likelihood remains that this method will be more suited to the literate, or to those more used to the manipulation of symbolic entities. It is difficult to see how this might be overcome, but unless there is evidence that ability to work in this way influences the type of sex that a person has, it is a relatively unimportant problem.

E. Invention

Inevitably, people are going to invent some of the entries in the diary. This may take the form of recording sex which they see as more acceptable, that is for example, leaving out the unsafe or the unsavoury, or either adding or omitting sessions so as to conform to a perceived notion of the 'right' amount of sexual activity.

It is assumed that when sessions are invented the invention will flow with the pattern of real sessions, so that if the focus of interest is the expectations of sexual activity or what I have referred to as the language of sex, then this is less of a problem than it would be if the focus was the incidence of sexual activity.

References


Appendix: Notes for a sexual diary (Project SIGMA)  

I. Introduction

Please read these Notes before starting your diary.

In your diary we want you to keep a daily record of your sexual activities, written in terms of what happened from your position. The Diary is to be written down on the weekly Diary Sheets as soon as you can after they occur and, if at all possible, on the same day. The accurate detail of your sexual behaviour is very important, and you must be completely honest or it isn't worth doing the diary. Although we are mainly concerned with homosexual behaviour, please make sure that you also record any sex you have with females during the month, describing it in the same detail.

You will probably find it useful first to get a sheet of paper and write down in your own words exactly what you can remember of the last time you had sex. Then read these Notes and check whether you got down all the significant information. You may write your diary either in ordinary words, or in a coded form (explained later), or in a mixture of the two. A word about language: in the Notes we shall usually refer to sexual activity using everyday terms, like "wank", rather than "masturbation", but in your Diary feel free to use whatever words you are comfortable with. Describe the sexual act as simply and accurately as you can, following the rules mentioned below.

(If you have a cassette recorder, and prefer to record rather than write down your sexual diary, this is quite alright by us; just make sure you give the same information and follow the same rules. Use your own cassette, and we shall return it to you after it has been transcribed).

The basic entry in the diary is what we call the "session". By "session" we mean: "one or more sexual acts with the same person (or persons) at one time". It might be a long session of different sexual acts, or simply be a quick wank by yourself. Within each session each "sexual act" (such as sucking off) is described separately, in terms of "who did what to whom, and who came".

Over the page we give an example of a day's entry in someone's sexual diary, consisting of one "session" at 8 am and another at 6 pm: it is written in ordinary language. Following this, the same day's entry is given, written in code.

---

2. Version of 17th November 1987

---

**Example of a Diary Entry:**

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Friday</strong></td>
<td>8 am</td>
<td>Had a wank by myself at home, reading porn, but didn't come</td>
</tr>
<tr>
<td><strong>Friday</strong></td>
<td>6 pm</td>
<td>my regular boyfriend (P1) and I had a session when he came home. First I sucked him (but he didn't come). Then he fucked me, using a condom and KY, and came in me. After that he wanked me off, using his spit.</td>
</tr>
<tr>
<td><strong>16 May</strong></td>
<td>During this, we used poppers</td>
<td></td>
</tr>
</tbody>
</table>

(You will find a longer example of a full week's diary on page 6). You might find it useful to adopt the coding scheme, which is like the one we use in the project. It looks a bit complicated, but is much shorter and actually very easy to get used too. Here is a coded version of the same diary entry:

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Friday</strong></td>
<td>8 am / home / SW</td>
<td></td>
</tr>
<tr>
<td><strong>16 May</strong></td>
<td>6 pm / P1's flat / P1 (AS + PFXOIC.1 + PWOKH1) / p</td>
<td></td>
</tr>
</tbody>
</table>

The Code and suggestions for its use are described in SECTION THREE (CODING SCHEME FOR SEXUAL ACTS) following these notes. You might like to look at it after reading through this first section. It contains a coded version of the same full week's diary.

Please do your best to keep your sexual diary on a daily basis, and keep it for the full month if you possibly can, because this is the shortest time in which it is possible to get an idea of your "typical" behaviour. If you cannot (or do not) complete a full month, return it just the same. Needless to say, if you are prepared to keep it either for a longer period than a month or every few months, this will be even more valuable. Mention this on the front sheet of the Diary.

**II. What needs to be described**

It is important that for each session you record:

- the time
- the place
- the participant/s (if any, other than yourself)
- what actually happened, including details of who came, and
- "accompaniments" (other things that are used or done, like using poppers).
  Let me explain each one in turn:

A. Time

Give the hour at which the session occurred if you can, or at least indicate whether it was morning, afternoon, evening or night.

B. Place

The reason for giving the place is that it is important to know where the sex took place - at your place or his, or somewhere else, such as outside (parks, heaths, toilets, etc.). Be honest!

C. Other person(s)

First of all, we do not want to know the names of your sexual partners, but we do ask you to describe each of them by:

i) whether you have sex with him on a regular, occasional or "one-off" (one-night stand) basis
ii) his age
iii) how long you've been having sex with him
iv) where you met him (if he's a "one-off" casual partner)
v) other details (job, where he comes from, what attracts you about him)
v) HIV status of your partner/s (if you know)

If you don't say otherwise, we assume your partner is male.

You write each week of the diary on a single page (but feel free to write on the back of the page, or design your own diary form if it doesn't suit you for any reason). You enter details about each sexual partner on the Partner List when they first appear in your diary, and then use that same number throughout the month.

D. The sexual acts

Your accurate description of the sexual acts which took place are the core of the sexual diary, so please pay a special attention to this part.

Each "session" consists of one or more sexual acts. For each sexual act we need to know:

- exactly what was done
- whether you did it by yourself; or you did it to him; or he did it to you; or you both it to each other at the same time
- whether it resulted in orgasm ("coming")
- anf, if so:

- who came
- whether the orgasm was "in" or "on" the person
- whether a condom was worn, poppers used, etc.

People sometimes withdraw their cock before coming, when sucking and fucking, as in example (5) below. Here, his partner pulled out before coming to orgasm, but then came on him. If this happens to you, or your partner does this with you, please be especially careful to mention it.

Here are some examples:

1. I wanked myself off
2. I sucked P1 off and he came in my mouth
3. P3 fucked me, but I didn't come
4. I sucked P2 and at the same time I wanked him off; he came. I didn't
5. He fucked me and came (on me, but not in me)
6. P2 sucked me, but he pulled out before he came. Then he took a shower and I sucked him and he came.
7. I rubbed myself on P1's stomach until I came
8. We wanked each other, but only he came.
9. He wanked off on my chest.
10. P1 screwed me and came inside me. P2 then rimmed me.
11. P1 rimmed me, then screwed me and came inside me.
12. P2 sucked me and came inside me, then he sucked me with a dildo. Then I sucked him with the same dildo.

Note in these examples:

- that sexual partners are referred to as "P1", "P2", etc.
- that "wanking / sucking off", or "wanking / sucking and coming" are taken to mean that the wanking / sucking resulted in orgasm.

E. "Other things" (condoms, lubricants, poppers, drugs and toys)

There are other things you may do when having sex which also need mentioning. Here are some common accompaniments:

i) Preparations. It's often important to know what happened before you had sex - whether you met at a club or in a cottage, or after being out drinking, for instance. But also, some people get themselves in the mood by having a joint, or giving themselves an enema (douche). Other people like to use porn magazines or videos, and so on. Mention any of these if they apply to you.

ii) Condoms and lubricants. It is especially important to know whether you or your partner/s use a condom, or any other form of protection. Mention the sexual act/s in which it was worn, and who wore it. Say, too, what type it was (e.g. Durex, Red Stripe) and what lubricants you use with it (e.g. KY,
121, Vaseline) at the bottom of the page. Sometimes you may begin to use a condom and then decide not to, or it might break or tear; mention this too.

For example:

He fucked me, wearing a condom.
P1 (wearing a condom) screwed P2, then P1 (wearing the same condom) screwed me.

Whilst fucking we tried to use a condom (Durex), and I had to use spit, but it broke in use after I came.

I wanked myself using baby oil as a lubricant.

ii) Poppers and drugs. If you ever use poppers (or any other drug/s) when having sex, mention this (and the name of the drug if it's not poppers). Once again, make it clear which sex act you use them in. For example:

First I wanked P1, then we had a "69" (mutual sucking), then I was fucked by P1 ... we both used poppers throughout.

He gave me poppers. First he rimmed me, and then he fucked me till he came.

After oiling each other down, we had a "69" (neither of us came).

Then we had a joint each, and then went on to...

iv) Toys, dildoes, etc. If you use any toys (such as a dildo, or equipment used in S&M), mention this, and when you use them. For example:

I was gagged and blindfolded, and then P1 ...
P2 sat on my stomach, and fucked me with my dildo, using KY.

To make it easier, there's a note at the bottom of each diary page where you can write in exactly what sort of condom, lubricant, poppers, toys etc. you use, or you can just put in when relevant.

Three (and more) some

When there are three or more people involved in a sexual session, things sometimes get more complicated. Always remember to say which person did what to whom. It's easiest to use "P1", "P2", "P3" etc to do this. You won't always know what the other people are doing, but if you do, mention it. For example:

I fucked P2, whilst he was fucking P3 at the same time

Whilst P3 and I were having a "69", P2 wanked off over me

P1 was sucking P2 whilst P2 was wanking me

---

Example of a full diary (Week beginning: Sunday 11 May)

<table>
<thead>
<tr>
<th>Day</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunday</td>
<td>9 am at my home, with P1. He fucked me (without coming), then I fucked him and came in. Both of us wore a condom (always Red Stripe) and used KY. Then I was wanked off by him.</td>
</tr>
<tr>
<td>May 11</td>
<td>11 at night, my place, reading porn, I had a wank by myself and came.</td>
</tr>
<tr>
<td>Monday</td>
<td></td>
</tr>
<tr>
<td>Tuesday</td>
<td>9 pm, Hampstead Heath. With P2 we both wanked each other off (at the same time) and Tuesday, both came. We shared poppers.</td>
</tr>
<tr>
<td>Wednesday</td>
<td></td>
</tr>
<tr>
<td>Thursday</td>
<td>7 pm at P1's flat. I sucked him, then he sucked me. Then I screwed him (using R/S &amp; Thursday KY). Then I was wanked off by him, and I then wanked him and he came.</td>
</tr>
<tr>
<td>Friday</td>
<td>7 am at home. A quick wank to orgasm. 11.30 pm, at my place, with P1. I was fucked by him and he came (but not in me). Then I fucked him without coming (KY but no condom in either case). Then I was wanked off.</td>
</tr>
<tr>
<td>Saturday</td>
<td>9 am, at home with P1: mutual wank, followed by me fucking him and wanking him at the same time (he came; I didn't). Then I wanked myself off. 4 pm, in this guy's (P3) car. I fucked him and came in him (with R/S &amp; KY), then I wanked him to orgasm. 11 pm, with P1 at my place. I wanked him, then he wanked me; neither of us came. Then we had a 69, and both of us came. We used poppers.</td>
</tr>
</tbody>
</table>

Type/s of:

CONDOMS used: always Red Stripe

LUBRICANTS used: KY

POPPERS used: Rush, Locker-Room

DRUGS used: (none)
FINALLY...

If you aren't having much sex at the moment, please don't invent sexual activity; your diary is just as important to the research as those who are very sexually active. Remember, too, that we are concerned with what you do, not with what you think you should do, nor even with what you would like to do! In these days, there's sometimes a tendency to describe our sexual behaviour as more safe than it really is. Please be completely honest.

Have a look now at the example of a week's diary which follows: then off you go! By the way, any resemblance between what's in the Diary and what you do is purely coincidental. The example has been written to give examples of different things that have been discussed earlier, and so there's lots more sexual activity in it than many people have in a month.

In section Two, which follows, there is a brief introduction to the Coding Scheme. You might like to have a look through it and then decide whether you want to write all your diary in it, or use bits of the code, or ignore it entirely. It's up to you.

MANY THANKS FOR YOUR UNIQUE AND VALUABLE INFORMATION. IT WILL BE KEPT ENTIRELY ANONYMOUS AND CONFIDENTIAL. Please write any additional information or explanation on the Diary forms, on the back of them, or on an enclosed piece of paper. Your thoughts, reactions and additional information are very valuable in understanding what is happening. Remember, too, that you may write to us (no stamp) at the Project at any time using the address given on the FREEPOST envelope.

Take care - and thanks again.

III. Coding scheme for sexual acts (Version of October 1st 1987)

Each sexual act is described basically in terms of what is done (type of act, e.g. sucking); which "way" it was done (i.e. by whom and to whom); whether and how orgasm occurred, and what accompaniments were used.

A. Type of act (what sexual act is done)

The first letters are used to stand for the type of sex act, as follows:

<table>
<thead>
<tr>
<th>Letter</th>
<th>Meaning</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>&quot;wanking&quot;</td>
<td>[i.e. masturbation]</td>
</tr>
<tr>
<td>S</td>
<td>&quot;sucking&quot;</td>
<td>[i.e. fellatio]</td>
</tr>
<tr>
<td>F</td>
<td>&quot;fucking&quot;</td>
<td>[i.e. anal intercourse]</td>
</tr>
</tbody>
</table>

A fuller list of sex acts is given on page 75.

B. Which "Way" (who does what to whom)

The "way" is coded by a letter put before the sexual act, to show how it was done.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Meaning</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>&quot;self&quot;</td>
<td>[i.e. you did it by yourself]</td>
</tr>
<tr>
<td>A</td>
<td>&quot;active&quot;</td>
<td>[i.e. you did it to him]</td>
</tr>
<tr>
<td>P</td>
<td>&quot;passive&quot;</td>
<td>[i.e. he did it to you]</td>
</tr>
<tr>
<td>M</td>
<td>&quot;mutual&quot;</td>
<td>[i.e. he did it to you and at the same time you did it to him]</td>
</tr>
</tbody>
</table>

So, for instance:

- SW means "Self-wank" i.e. you wanked yourself
- AS means "Active-suck" i.e. you sucked his cock
- PF means "Passive-fucked" i.e. you were fucked by him
- MS means "Mutual-sucked" i.e. you sucked each other's cock at the same time ("69")

C. Orgasm

It's important to be very specific about orgasm. In particular, we need to know:

- if anyone came to orgasm. And if so, who did?
- whether the orgasm was "in" or "on" the other person, and
- if "in", whether a condom was worn,
- what other accompaniments were used.

If no orgasm occurred, then nothing needs to be added to the sexual act, so that, for instance, AF simply means you fucked him, and neither of you came to orgasm in that act.

If orgasm did occur, the details are coded by adding two letters to the end of each sexual act: this first to describe your orgasm, and second to describe his orgasm. To describe the obvious alternatives in each position, we use three different letters:

<table>
<thead>
<tr>
<th>Letter</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>is used to mean you (or he) came to orgasm</td>
</tr>
<tr>
<td>X</td>
<td>is used to mean you (or he) did not come at all</td>
</tr>
<tr>
<td>P</td>
<td>is used to mean you (or he) pulled out before coming, i.e. came &quot;on&quot; rather than &quot;in&quot;.</td>
</tr>
</tbody>
</table>

If the sex act is "single", i.e. only you were involved, then there is no need to include the second position, and only O applies, as in:

- SWO means "single-wank-orgasm" [i.e. you wanked yourself to orgasm].
B. Which "Way" (who does what to whom)

The "way" is coded by a letter put before the sexual act, to show how it was done.

S means "self" [i.e. you did it by yourself]
A means "active" [i.e. you did it to him]
P means "passive" [i.e. he did it to you]
M means "mutual" [i.e. he did it to you and at the same time you did it to him]

So, for instance:

SW means "Self-wank" i.e. you wanked yourself
AS means "Active-suck" i.e. you sucked his cock
PF means "Passive-fucked" i.e. you were fucked by him
MS means "Mutual-sucked" i.e. you sucked each other's cock at the same time ("69")

C. Orgasm

It's important to be very specific about orgasm. In particular, we need to know:

- if anyone came to orgasm. And if so, who did;
- whether the orgasm was "in" or "on" the other person, and
- if "in", whether a condom was worn,
- what other accompaniments were used.

If no orgasm occurred, then nothing needs to be added to the sexual act, so that, for instance, AF simply means you fucked him, and neither of you came to orgasm in that act.

If orgasm did occur, the details are coded by adding two letters to the end of each sexual act: this first to describe your orgasm, and the second to describe his orgasm. To describe the obvious alternatives in each position, we use three different letters:

O is used to mean you (or he) came to orgasm
X is used to mean you (or he) did not come at all
P is used to mean you (or he) pulled out before coming, i.e. came "on" rather than "in".

If the sex act is "single", i.e. only you were involved, then there is no need to include the second position, and only O applies, as in:

SWO means "single-wank-orgasm" (i.e. you wanked yourself to orgasm).

FINALLY...

If you aren't having much sex at the moment, please don't invent sexual activity; your diary is just as important to the research as those who are very sexually active. Remember, too, that we are concerned with what you do, not with what you think you should do, nor even with what you would like to do! In these days, there's sometimes a tendency to describe our sexual behaviour as more safe than it really is. Please be completely honest.

Have a look now at the example of a week's diary which follows: then off you go! By the way, any resemblance between what's in the Diary and what you do is pure coincidence. The example has been written to give examples of different things that have been discussed earlier, and so there's lots more sexual activity in it than many people have in a month.

In section Two, which follows, there is a brief introduction to the Coding Scheme. You might like to have a look through it, and then decide whether you want to write all your diary in it, or use bits of the code, or ignore it entirely. It's up to you.

MANY THANKS FOR YOUR UNIQUE AND VALUABLE INFORMATION. IT WILL BE KEPT ENTIRELY ANONYMUS AND CONFIDENTIAL. Please write any additional information or explanation on the Diary forms, on the back of them, or on an enclosed piece of paper. Your thoughts, reactions and additional information are very valuable in understanding what is happening. Remember, too, that you may write to us (no stamp) at the Project at any time using the address given on the FREEPOST envelope.

Take care - and thanks again.

III. Coding scheme for sexual acts

(Version of October 1st 1987)

Each sexual act is described basically in terms of what is done (type of act e.g. sucking); which "way" it was done (i.e. by whom and to whom); whether and how orgasm occurred, and what accompaniments were used.

A. Type of act (what sexual act is done)

The first letters are used to stand for the type of sex act, as follows:

W means "wanking" [i.e. masturbation]
S means "sucking" [i.e. fellatio]
F means "fucking" [i.e. anal intercourse]

A fuller list of sex acts is given on page 75.
If the act is Active, Passive or Mutual, add two letters, for "ME" and "HIM", and fill in the appropriate position. First, using O for "Orgasm" and X for "Didn't come":

<table>
<thead>
<tr>
<th>(Me)</th>
<th>(Him)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only I came to orgasm:</td>
<td>0</td>
</tr>
<tr>
<td>Only he came to orgasm:</td>
<td>X</td>
</tr>
<tr>
<td>Both he and I came to orgasm:</td>
<td>O</td>
</tr>
</tbody>
</table>

Examples:
- MW OX means "mutual-wank; I came, he didn't"
- MS XO means "mutual-suck; I didn't come, he did"
- MW OO means "mutual-wank; we both came"
- AF OX means "active-fuck; I came, he didn't"
- AF XO means "active-fuck; I didn't come, but he did" (i.e. he came as a direct result of me fucking him)

Now, using P, as well as O
- MWPIX means "mutual-wank; I pulled out before coming, and he didn't come at all"
- MSOP means "mutual-suck; I came (in his mouth) but he pulled out before coming"
- AFOO means "active-fuck; I came in him and he came too"

D. Other

The most important "accompaniments" of a sexual act are condoms, lubricants, poppers and/or drugs and toys. These "other bits" are coded after the sexual act, using a slash (/), with the appropriate small (lowercase) letters, so:

<table>
<thead>
<tr>
<th>add</th>
<th>for condom</th>
</tr>
</thead>
<tbody>
<tr>
<td>add</td>
<td>for lubricants</td>
</tr>
<tr>
<td>add</td>
<td>for poppers (or the name of the drug)</td>
</tr>
<tr>
<td>add</td>
<td>for dildo (inserted anally)</td>
</tr>
</tbody>
</table>

If they are used in the act. For instance:
- AF/c.l means "Active-fuck / condom and lubricant"
- AS/p means "Active-suck / poppers"; i.e. whilst you sucked him, poppers were used.

Coding a session

If more than one sexual act was performed in a "session": put the various acts together (with a "\+" between each) in the order in which they occurred, then put brackets round the session. For instance:

(PS + AFOX/c.l + PP) / p

In words: you were sucked by him, and then you sucked him and came (you were wearing a condom for that act). Then you were fucked by him (but neither of you came). During the whole session, poppers were being used.

If this coded version all seems too complicated, just use your own words - but be sure to mention the things mentioned here. You may, if you wish, mix the "full form" of reporting your diary and the coded form. In this case, use the codes for the main description, and put in full any bits you're unsure about. A coded version of the week-long diary, a Reference Guide to Coding and on APPENDIX (giving details about encoding complex acts and 3-somes) follows.

Example of a (coded) week's diary
(Week beginning: Sunday 11 May)

<table>
<thead>
<tr>
<th>Day</th>
<th>Time</th>
<th>Location 1</th>
<th>Location 2</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunday</td>
<td>9 am / home</td>
<td>P1 (PF / c.l + AFOX / c.l + PWOX)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 11</td>
<td>11 pm / home</td>
<td>SWO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monday</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tuesday</td>
<td>9 pm / park</td>
<td>P2 (MWOO)</td>
<td>/ p</td>
<td></td>
</tr>
<tr>
<td>Wednesday</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thursday</td>
<td>7 pm / his flat</td>
<td>P1 (AS + PS + AFOX / c.l + PWOX + AWXO)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friday</td>
<td>7 am / home</td>
<td>P1 (PFXP / AF + PW + PWOX)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturday</td>
<td>9 am / home</td>
<td>P1 (MW + AF &amp; AW XO + SWO)</td>
<td>4 pm / car</td>
<td>P3 (AF / c.l + AWXO)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11 pm / P1 (AW + PW + MSO)</td>
<td>/ p</td>
</tr>
</tbody>
</table>

Type/s of:
- CONDOMS used: always Red Stripe
- LUBRICANTS used: KY
- POPPERS used: Rush, Locker-Room
- DRUGS used: (none)

73
IV. List and code of sex acts

The Code for each act consists of three parts:
- how it was done (Self, Active, Passive, Mutual)
- what was done (Wanking, Sucking, Fucking, etc)
- whether and how orgasm occurred, for me and for him, and what accompaniments were used.

A. List of more common sexual acts

[WANKING]

SW  Self Wanking
AW  Active (you did it to him) Wanking
PW  Passive (it was done to you by him) Wanking
MW  Mutual Wanking (you both did it to each other at the same time)

[SUCKING]

AS  Active Sucking
PS  Passive Sucking
MS  Mutual Sucking ("69")

[FUCKING (anal intercourse)]

AF  Active Fucking
PF  Passive Fucking
MF  Mutual Fucking

[FUCKING (vaginal intercourse with a woman)]

AV  Active (vaginal) Fucking

Note: Several other variations are possible, but uncommon, like SS (self-sucking) and SF (self-fucking): if you do it, put it down in code!

[OTHER SEXUAL ACTS]

Two letters are used to stand for the other sexual acts:

BR  Body Rubbing (face-to-face, as in "Princeton Rub")
CP  Corporal Punishment (belt, whip, hands)
FG  Fingering (insertion of finger into anus)

FI  Fisting (anal insertion beyond knuckle)
MA  Massage
RI  Rimming (oral-anal; need not be insertion)
SC  Scat (shitting)
TF  Thigh Fucking (cock between his thighs / legs / armpits, etc.)
WS  Water-sports (pissing)

(Invent your own 2-letter symbol for any acts not included here, but explain them). As in the previous set of acts, put an "S" before the act for "Self", put an "A" before the act for "Active" and "P" for "Passive", and put an "M" before the act for "Mutual". For example:

A  MA  you massaged him
P  TF  you were thigh-fucked by him
M  RI  you both rimmed each other at the same time

Two (or more) different acts which occurred at the same time are joined by 
"&" (AND) (e.g. AF & AW), and three-somes are encoded using Partner numbers (see APPENDIX at the end).

B. Coding orgasm

Add to the relevant sex act two letters representing the orgasm of ME and HIM [or, for 3-somes, for the two partners involved, e.g. (P2, P3)]

(Me)  (Him)

| Only I came to orgasm: | 0   | X   |
| Only he came to orgasm:| X   | O   |
| Both he and I came to orgasm: | O   | O   |

In the case of Single sex acts, only one letter need be added, as in:

SWO  meaning "I wanked myself to orgasm"

Instead of O, the following variants should be used when appropriate:

X  is used to mean you or he did not come at all
P  is used to mean you or he pulled out before coming, i.e. came "on" rather than "in"

C. Accompaniments

Add to the relevant sexual act or session:

κ  means a condom was used
λ  means a lubricant was used
Al  means a dildo was used
D. Other

Spaces and brackets can be used whenever necessary to make clear what happened (e.g., whether poppers were used throughout a session or only in a single act).

V. Note on coding simultaneous acts and threesomes

A. Two different acts at the same time

Sometimes two (or more) different sexual acts are done together at the same time. If so, link them by an "&", meaning "AND". They then count as a single act. Thus:

- PS&AW means "Passive-suck-AND-Active-wank"
  i.e. he sucked you and you wanked him at the same time

- AF&AW means "Active-fuck-AND-Active-wank"
  i.e. you fucked him and you wanked him at the same time

Mutual sexual acts are simply the same sex act done at the same time by two people, and sometimes it might help to use the "&" to make clear exactly what happened. Consider:

- AF&PF which means "Active-fuck-AND-passive-fuck"
  i.e. you fucked him and you were fucked by him at the same time

We believe (so far!) that this sexual act is physically impossible so MF can't be used to describe it. But it could have been a threesome, where there would be two partners involved with you at the same time. If this is so, make it clear by adding their numbers, e.g.,

- AF(P1)&PF(P2) means "Active-fuck (by P1) - AND - passive-fuck (by P2)" i.e. in a 3-some you fucked P1 and at the same time, you were fucked by P2

A quite different act is coded by PF(P1)&PF(P2), which is possible, if unusual:

- PF(P1)&PF(P2) means "Passive-fuck (by P1)-AND-passive-fuck (by P2) i.e. you were fucked by P1 and you were fucked by P2 at the same time.

When orgasm occurs, simultaneous acts are encoded in exactly the same way, so:

(AF&AW) XO means "(active-fuck-AND-active-wank), I came, he didn't"
[in this case I was fucking him and wanking him, and it was as a result of wanking him that he came.]

B. Threesomes

In the case of three-somes when orgasm/s occur, it is sometimes difficult to make clear exactly who came, and here it sometimes becomes necessary to include the orgasm information with each constituent act, and specify the partner. Consider the following three-some involving me, P1 and P2:

PF XO (P1) & PS OX (P2) & AW (P1,P2) XO
i.e. I am fucked by P1 (and P1 comes) AND at the same time I am sucked by P2 (and I come) AND at the same time P1 wanks P2 (and P2 comes).

The only new rule is using (P1, P2) to show who the orgasm information refers to when it doesn't include me, so AW (P1, P2) XO translates as "Active-wank (involving P1, P2); P1 didn't come, P2 came".

We recognize that a code such as this cannot be complete, but the applications of these rules mean that almost every session can be encoded. Clearly, if you are in a three-some you will not necessarily know what the others are doing, but if you encounter a situation which you think cannot be encoded, please inform us.
Research Agenda

THE STRUCTURE OF SEXUAL BEHAVIOUR

ANTHONY P. M. COXON, Ph.D., PETER M. DAVIES, Ph.D.,
University of Essex
ANDREW J. HUNT, B.A., PETER WEATHERBURN, M.Sc.
South Bank Polytechnic
THOMAS J. McMANUS
King's College Hospital
CLIVE REES, B.Sc.
University of Wales

In a national study of gay men's sexual behaviour in the context of HIV transmission, a schema was evolved for encoding and reporting the structure of sexual activity: the Sexual Behaviour Code (SBC). It has linguistic properties, is easily learned and employed, and generalizes easily to heterosexual behaviour. Each sexual session (sentence) comprises a sequence of sexual acts (words), which in turn consists of a sexual behaviour (root), the modality (prefix), and the outcome/ejaculation (suffix). Other aspects (partners, drugs, condoms) are encoded as precedents and accompaniments of the acts. The SBC is also used to define a comparable and flexible research instrument for eliciting systematic information on sexual behaviour: the Inventory of Sexual Behaviour (ISB), now used in all WHO/GPA studies of homosexual response. Issues of validity and reliability are addressed, and forms of analysis of the resulting data are discussed for sexual diaries and inventories.

KEY WORDS: Sexual behaviour, sexual diaries, sexual inventories, homosexual, HIV transmission

The need to develop a schema for describing sexual behaviour arose initially in the attempt to elicit systematically the full details and variants of sexual activity which might conceivably be implicated in the transmission of the Human Immuno-deficiency Virus (HIV) among gay and bisexual men, since this was the research remit of Project SIGMA.¹ A review of literature describing such behaviour revealed a wide range of schemes, some of which were incomplete (e.g., not enquiring about common behaviours such as oral-anal contact ('rimming'))²; others were inconsistent, and yet others used categories which were

¹The acronym stands for Socio-sexual Investigations of Gay Men and AIDS. The research in this article was funded by the Department of Health (UK). The views expressed in this article are those of the author and not necessarily those of the Department.
²In this paper, commonly used terminology will be used whenever a subject's own words are cited, or when the neutral medical terminology would be cumbersome.

Requests for reprints should be sent to Dr. Anthony P. M. Coxon, Department of Sociology, University of Essex, Wivenhoe Park, Colchester Essex CO4 3SQ, England.
not sufficiently specified to differentiate between two similar varieties of behaviour which had differing consequences for the probability of HIV transmission.3

Undoubtedly, the hitherto most extensive schema for describing sexual behaviour was that developed by the Kinsey Institute, both in the original study (Kinsey et al., 1948) and in the study of homosexuality (Bell & Weinberg, 1978). Their method of coding and shorthand referencing provided a basis for our own.

The second motivation was the need to develop adequate and compatible "reporting structures" for relating different national studies, and later to produce a common protocol for research studies under the auspices of the World Health Organization.4 Once again, a survey of the literature revealed very great differences in question wordings and reports, a bewildering range of types of response permitted or provided, and grossly incompatible differences in the time-span of reporting (or, indeed, the failure to specify any time span at all).5

For national studies covering very different cultural manifestations of homosexual behavior, it was critical to ensure as a minimum requirement common reference to the physical aspects of sexual behavior.

The initial research context in which the schema was developed was experimentation with the method of sexual diaries (Coxon, 1988a, 1990; Davies, 1989; Davies & Coxon, 1990). Presently, the schema is used as a primary data eliciting tool by Project SIGMA and has formed the basis for developing an "Inventory of Sexual Behaviour" for use in interviews and questionnaires. The schema also yields exactly compatible data at any chosen level of aggregation.

Structure of the Sexual Act

As a first step, we restricted attention to the behavioral aspects of sexual activity. We also made a simplifying assumption: that sexual behaviour is (or, more accurately, can be described as) a discrete rather than a continuous phenomenon. This discreteness is an over-simplification, since physical sexual interaction manifests continuous (if not smooth) movement. On the other hand, it is virtually impossible to describe such behaviour in a genuinely continuous manner, and in practice people experience the continuity as "chunked" into more or less discrete actions which can be described verbally.

6Such as the use of the single generic term "anal intercourse" to cover both receptive and active modes.

7The initial national studies consisted of the Netherlands Study (Department of Homosexuals, University of Utrecht, Dr. R. A. P. Telman), the San Francisco Study (Center for AIDS Prevention Studies, Dr. Tom J. Coates and Dr. Susan Kegels) and Project SIGMA (Drs. P. M. Davies and T. J. McKinnon, London, and Dr. A. P. M. Coxon, Cardiff, Wales). The WHO research projects were organized as the Homosexual Response Studies under the direction of Dr. Manuel Carballo (WHO Global Program on AIDS, Social and Behavioural Research) and the convenorship of the author.

8Some of these problems are discussed in Coxon (1990b) and an attempt is made there to render some schemes consistent in order to produce enough compatibility for relative correlation.

9This is by no means a trivial issue. In training researchers to code, we used a sexually explicit video as a common reference stimulus and compared researchers' "scripts" after a short training session. Although there was agreement over gross characteristics (such as the fact that solo masturbation occurred), there were considerable differences over the level of fine-ness of description. Some tended to lump actions together, whilst others made finer distinctions, splitting actions into separate units. Nonetheless, these seemed to be an

THE STRUCTURE OF SEXUAL BEHAVIOUR

Content of Researchable Aspects of the Sexual Act

If such a schema is to be generally useful, it must be capable of being used in a variety of data-eliciting situations. In particular, it must be applicable when the subject makes his own descriptions (self-reports); when the sexual activity is observed (and/or participated in); and in the interview or self-completed questionnaire. It must also be easily comprehensible (if subjects are going to be able to use it), and it should be encryptable, so that secrecy and confidentiality can be assured. Finally, a more important requirement is that the schema must be detailed enough and also "open" to modification; it must be complex enough to encompass all common (and most rare) sexual behaviour and situations, and allow the addition of new behaviour as it occurs.

Restrictions on Manageability

The schema we have developed is dubbed the "Sexual Behaviour Code" (SBC). Certain initial restrictions on coverage are made, not out of principle but for practical reasons. First, in terms of the structure of the sexual act, we are concerned only with binary (or unary) behavior, in the sense that the description is defined from ego's (the participant's) perspective: what is done to him or by him. If there are more than two participants, then any activity involving ego is relevant, but not (for example in a three-person) all activity between the other participants.8

Although the schema presented here was developed to describe and encode male homosexual behaviour, there is no reason why it cannot be extended to describe any sexual activity, and it is a trivial matter to extend it to heterosexual behaviour.9 Furthermore, the schema has also been implemented in two quite different methodological forms: the Sexual Diary10 and the Inventory of Sexual Behaviour.11

acceptable common level of description, corresponding in many ways to the "natural level" or "generic taxa" found to characterize many folk taxonomies in cognitive anthropology (Berlin, 1978), and these differences in behaviour "virtually 'cry out to be named,"' as Berlin graphically puts it (op cit, p. 24). We are grateful to IIM Customs and Excise who have allowed us to access such video material for research purposes before returning them to classified status.

9With the advent of "safer sex" and the proscription of some activities such as anal intercourse, gay men have become more inventive and developed various alternatives, such as oral-genital and oral-anal penetration ("kissing-fucking").

10Another version of the SBC is the n-ary description. This is defined from the observer's perspective, and each participant needs separate identification in the code. In this context, different rules are needed for combination and constitution of sexual relationships. Davies (1989) has devised such an n-ary system, using video sequences as sources.

11The modifications involve specifying the sex of the participant and the extension of sexual behaviours to include exclusively heterosexual variants such as vaginal intercourse.

The method of sexual diaries used in these studies is documented in Coxon (1988a, 1990b) and further material (including detailed instructions to respondents) is contained in Notes and Revised Notes for the Sexual Diary (Coxon, 1986 and Coxon, 1986c, respectively).

12The Inventory exists in various forms, containing differing degrees of detail. The Basic Form (as used in the WHO/GPA Homosexual Response Studies, Core Questionnaire) restricts attention to the most common sexual acts and asks only whether each variant has ever (in the last 6 months) been performed. The Extended Form (HIS Non-Core) contains additional acts and asks for specification by each partner.
Common Structure of Sexual Behaviour

A diagram of the basic structure of sexual behaviour (in our account) is given in Figure 1.

![Figure 1 Components of the Sexual Session](image)

THE SEXUAL SESSION: FRAMEWORK OF SEXUAL ACTIVITY

The main unit of sexual behaviour is The Sexual Session, denoted by the outer box in Figure 1. Using a linguistic analogy, which we shall later exploit more literally, the sexual session may be considered as the "sentence" of sexual activity—self-sufficient and intrinsically well-formed.

The sexual session occurs at a given (specified) time and place, consists of one or more sexual acts, and involves at least one, usually two (and sometimes more) people. It is typically terminated by sleep, a nonsexual intermission or change of partner. The characteristics of a Sexual Session thus include four components: the Setting, the Precedents, the Accompaniments, and the Partner Specification. The Setting refers primarily to when and where the sexual activity took place. "When" is usually the time of day, so that sexual sessions can be kept in time order, and the "Where" usually refers to the location (such as "boyfriend's flat" or "my home", or outside locations such as "Public toilet at X", or "Park Y"). The Precedents refer to any relevant events which preceded or led up to the sex, such as preludatory drinks or drugs, use of poppers (nitrates) or stimulants such as video or magazines. The Accompaniments refer to precedents which continued during the sexual activity, together with the use of additional objects during the sexual activity, such as lubricants and "toys" (e.g., dildoes, leather or other costume). And finally, Partner Specification would ideally consist of the name and characteristics of the partner, such as age, sex, occupation, and relationship status. However, it is often necessary to remove the actual name of a partner.

The characteristics requested depend on the purpose of the analysis. In some cases it is feasible to ask the identity of the partner. Where this is not feasible (for example, in Project SIGMA), each partner is assigned a unique (arbitrary, but sequential) number and that is how he be described in terms of Partner Status (Regular, Occasional, Casual, One-off). Age, How long the subject has been having sex with the partner, Where the partner was met on this occasion, HIV antibody status (if known), and "Other" information, if known, such as occupation and characteristics found especially noteworthy, attractive or repulsive and whether payment was involved.

from the record when anonymity or confidentiality has been given, but in so doing all record linkage (for network or contact analysis) is then forgone.

The core component in describing sexual activity is what actually happens—in our terminology, the "sexual act."

The Sexual Act: Behaviour, Modality and Outcome

On the linguistic analogy, the sexual act is the word in the sentence, and sexual acts make up the sexual session. It is the sexual act which specifies "WHO does WHAT and with what EFFECT." These three components are called: (1) the Behaviour (or behaviours) which refers to the actual sexual activity itself ("what" is done); (2) the Modality which refers to "who (does the activity), and to whom"; and (3) the Outcome which refers to the "effect" of this sexual activity, which in the context of HIV transmission becomes the question of whether ejaculation of sperm occurs, and if so, to whom and in what manner. Let us specify each of these in turn.

The Sexual Act: Behaviours

We assume that a succession of (continuous) bodily movements are, or can be, "chunked" into identifiable and (well-nigh) universally recognizable sexual activities and given a common name. For instance, whilst the act of masturbation will usually have at least some unique components (for no-one does it exactly the same way, and no-one repeats the act identically), the manual stimulation of the penis by the hand is usually taken to be a necessary part of the definition. Thus, random, unintended rubbing of the penis would be excluded from the category of "masturbation."

Despite its virtual universality, and the early age at its first appearance, the nomenclature for masturbation is far from universal. Because the act of masturbation is often taboo especially among children, and frequently discouraged or prevented, it comes to be referred to by all sorts of euphemisms and code-names (often unique to the family concerned). There thus arises a hierarchy of terms of differing acceptability, from the medical terminology used professionally (and often when talking to professionals) through a widely-used set of vernacular terms to largely idiosyncratic ones. In Project SIGMA, the WHO studies, and in related studies, we have always elicited such terminology before proceeding to questioning on detail of sexual behaviour. The purpose of this is not only to gather information on "street" terminology but also to make the respondent more at ease in asking detailed information about what may be an embarrassing topic. For this reason, all our research forms including the Inventory of Sexual Behaviour are constructed with the neutral (medical) term.

13 See Kinsey et al., 1948, pp. 497-498; "the word 'masturbation' refers to any self-stimulation which is deliberate and designed [sic] to effect erotic arousal. By such a definition, the accidental touching of oneself is not masturbation because it is not deliberate," ibid., loc. cit.

14 As well as being a device for establishing preferred terminology for the interviews, it also serves as a means of collecting systematic data on the distribution of such terms and on the contexts of their use. For instance, it happened not infrequently that a respondent would give "wank" as his usual word among friends, but prefer to use "masturbation" in the interview.
in curly brackets, and the subject's chosen alternative is substituted in the
spoken form. Thus:

"Have you ever masturbated another man to his ejaculation?" may be
rendered in the interview as:

"Have you ever 'wanked' a guy until he came?"

The set of sexual acts must be open-ended. Although the number of distinct
and distinguishable sexual activities is immense, the distribution of types of
sexual acts is very skewed, and a surprisingly small handful of acts (typically
three) normally suffices to encompass over 95% of sexual expression. This is
well illustrated from our 1986 study of British gay men's sexual behaviour using
sexual diaries, where 95% of the total number of acts (not persons performing
the acts) were either masturbation, fellatio or anal intercourse (Coxon, 1990b).

But it is the less common behaviours which often occasion the most research
interest, for even their approximate prevalence is not known and yet their
risk-status for transmission may be very large. This is especially true of
ano-rectal (or Brachio-proctie) insertion ("fist-fucking"), which can be espe-
cially traumatic and—especially in conjunction with receptive anal intercourse
—implicated as a high-risk behaviour for the transmission of HIV. Project
SIGMA records 16% of British gay men as having ever engaged in it (Coxon,
1990b, p. 126), with an incidence of between 3% and 5% for sexual diary monthly
data (Coxon, 1989).

It is not usually difficult to specify the list of the more common or well-
established sex acts such as masturbation and fellatio; it is the more recently
developed and ill-defined ones which cause difficulty. Thus, "Massage" can mean
anything from rubbing hands over the body to systematic manipulation of
muscles, and a practice like inter-femoral penile insertion ("thigh-fucking"),
which tends to substitute for anal intercourse in a post-Aids era, can cover a
wide range of actual practices.

In the case of Project SIGMA and the WHO Homosexual Response Studies,
the following list is used.16 (The most common activities are encoded by a single
letter (W,S,F) and the rest by a two-letter code. Obviously, these letter allocations
are to some extent arbitrary, but in this scheme the Code letters refer to the
most common "street term" for the practice, thus making it easier for subjects
to make use of the code.)

<table>
<thead>
<tr>
<th>Behaviour</th>
<th>Code</th>
<th>Percent Of Act</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osculation (&quot;deep kissing&quot;)</td>
<td>Dk</td>
<td>——</td>
</tr>
<tr>
<td>Masturbation (&quot;wanking&quot;)</td>
<td>W</td>
<td>62%</td>
</tr>
<tr>
<td>Fellatio (&quot;lapping&quot;)</td>
<td>S</td>
<td>22</td>
</tr>
<tr>
<td>Anal intercourse (&quot;fucking&quot;)</td>
<td>F</td>
<td>9</td>
</tr>
<tr>
<td>Analusius (&quot;rimming&quot;)</td>
<td>Ri</td>
<td>3</td>
</tr>
<tr>
<td>Body Massage</td>
<td>Ma</td>
<td>1</td>
</tr>
<tr>
<td>Digital-anal insertion (&quot;fingerinsertion&quot;)</td>
<td>Fg</td>
<td></td>
</tr>
<tr>
<td>Ano-rectal insertion (&quot;fisting&quot;)</td>
<td>Fr</td>
<td></td>
</tr>
<tr>
<td>Corporal punishment</td>
<td>Cp</td>
<td></td>
</tr>
<tr>
<td>Interfemoral penile insertion (&quot;thigh-fucking&quot;)</td>
<td>Tf</td>
<td></td>
</tr>
<tr>
<td>Limnism (&quot;water-sports&quot;)</td>
<td>Wa</td>
<td></td>
</tr>
<tr>
<td>Nipple tweaking/nibbling</td>
<td>Tx</td>
<td></td>
</tr>
</tbody>
</table>

(* denotes < 1%)

It is a moot point whether (apart from vaginal intercourse) there is any sexual act which
does not apply in both homosexual and heterosexual behaviour, since knowledge of the
sex of the participants provides the necessary context for disambiguation (but are fellatio
and cunnilingus to be distinguished as separate acts?). However, additions which can
(and in the case of Project SIGMA Heterosexual Panel have been) added for male-female
sexual behaviour include:

12. Vaginal intercourse

13. Cunnilingus

(Cn)

Modality of the Sexual Act

The modality of the sexual act—the prefix to the sexual act "word"—specifies
which actor (from ego's standpoint) did what sexual act to which actor. A variety
of contrast names exists to distinguish these two roles: inserter/insertee,
active/passive, as well as "street" language such as the "door" (usually, inserter),
butch/bitch etc. Although particularly relevant to anal and vaginal intercourse,
the inserter/insertee contrast cannot be used generally since it carries with it a view
of sexual activity as primarily insertive (and by implication, male). More
importantly, the role-difference in many acts—such as masturbation—is not,
and cannot reasonably be described as, that of inserter/insertee. Therefore, we adopt
the "active/passive" distinction as basic, and use it in accord with conventional
grammatical usage to denote verbal mood (hence "modality"), noting that this
also conforms to common usage among gay men:

ACTIVE means that EGO does the given sexual act TO ALTER
PASSIVE means that the sexual act IS DONE TO EGO BY ALTER.

Thus, "active masturbation" means that I masturbate the other person
and "passive masturbation" means that the other person masturbates me. That the
active/passive distinction is not identical to the inserter/insertee distinction can be
seen by the following example:

Most gay men describe fellatio as "sucking," in which case "Active sucking"
means I insert my penis in the other person's mouth (i.e., I am inserter) and
"Passive sucking" means the other person inserts his penis in my mouth (i.e., I
am insertee). But another term for fellatio, used especially in the U.S.A. to stress
the "butch" aspect, is "mouth-fuck." In this case, "Active mouth-fuck" means I
am inserter, and the passive variant means I am insertee—thus reversing the
modality compared to "suck." Hence, the active/passive distinction is reliant on

---

16The percentages refer to the number of times the behaviour occurred in a month,
produced from the May 1986 SIGMA Sexual Diaries (N=171) data (Coxon, 1990b).
the meaning of the sexual act concerned, but (where relevant) it can simply be converted into the inserter/ee distinction.

But "active" and "passive" do not exhaust the alternatives of modality. A useful (but strictly speaking redundant) modality is mutuality, which means that both persons simultaneously do the same sexual act to each other. (This is not to be confused with the sequential performance of the same act, which is quite distinct).

Mutual means that Ego does the sexual act to Alter at the same time as Alter does the same act to Ego.

So far, all the modalities have been binary—relating two distinct persons. But for many sexual activities, there is only one person involved, usually ego. So we have to make allowance for "Self"—self-masturbation being the most common and obvious one. Logically, one can make allowance for the self-modality for any sexual act, but some are more likely (or indeed physically feasible) than others. Sufficient to say that we have encountered many subjects who digitally enter their own anus, some who are able to fiddle themselves and even a very occasional instance of penis insertion in one's own anus.

There is a further unary modality: when the sexual act is done by alter to him/herself. It might be thought that it would be excluded as irrelevant from an ego-centric account such as this. We were forced to reconsider this when we encountered instances of the other person doing something that could (at least in principle) lead to transmission of HIV to the subject—such as when alter masturbates and ejaculates onto the subject's chest, or uses his ejaculate as a lubricant for masturbating the subject—two relatively common activities which could lead to transmission of HIV if there were lesions on the recipient's chest or penis.

The various modalities are summarized in the following Table.

<table>
<thead>
<tr>
<th>Modality</th>
<th>Logical Character</th>
<th>In Words</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>Unary, reflexive</td>
<td>I do it to myself</td>
<td>S</td>
</tr>
<tr>
<td>Him/her</td>
<td>Unary, reflexive</td>
<td>She does it to him/herself</td>
<td>H</td>
</tr>
<tr>
<td>Active</td>
<td>Binary, asymmetric</td>
<td>I do it to him/her</td>
<td>A</td>
</tr>
<tr>
<td>Passive</td>
<td>Binary, asymmetric</td>
<td>She does it to me</td>
<td>P</td>
</tr>
<tr>
<td>Mutual</td>
<td>Binary, symmetric (equivalence)</td>
<td>We both do it to each other at the same time</td>
<td>M</td>
</tr>
</tbody>
</table>

The Outcome of the Sexual Act: Ejaculation

The dominance of HIV transmission in our account naturally means that we concentrate primarily on ejaculation of sperm as the primary "outcome" of the sexual act. Sometimes, this was euphemistically called "orgasm" and, indeed, there were those who said they had had orgasm without ejaculation, and at least one who claimed ejaculation without orgasm.

Initially, we concentrated simply on whether a subject ejaculated or did not as the result of a given sexual act. As time went on, the inventiveness of some members of the gay community meant that we had to pay more careful attention to exactly where the ejaculate went (perhaps under the Safer Sex guidelines which promulgated "On and not in your partner"), and by the increasing need to know whether that destination was a condom. Indeed, condoms began as "accompaniments" in our account and finished as an integral part of the outcome.

In explaining this issue, I shall at this point ignore outcomes other than ejaculation vs non-ejaculation (ejaculation will be referred to hereafter as "orgasm") and return in the subsequent section to condoms and other destinations.

For any (binary) sexual act, there are basically only four simple outcomes we wish to distinguish:

<table>
<thead>
<tr>
<th>Ejaculation:</th>
<th>Ego</th>
<th>Alter</th>
<th>In Words</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. No</td>
<td>No</td>
<td>Neither ejaculates</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2. No</td>
<td>Yes</td>
<td>Only he ejaculates</td>
<td>X 0</td>
<td></td>
</tr>
<tr>
<td>3. Yes</td>
<td>No</td>
<td>Only I ejaculate</td>
<td>O X</td>
<td></td>
</tr>
<tr>
<td>4. Yes</td>
<td>Yes</td>
<td>Both ejaculate</td>
<td>O</td>
<td></td>
</tr>
</tbody>
</table>

It might be thought (as indeed we originally did) that only a single outcome need be specified either because the act is unary (and only one person is involved) or because the combination of a given act with a particular modality leaves no doubt. Thus, if ejaculation occurs during Passive Anal intercourse, it must surely be my partner who ejaculates in me (at least as a direct result of the act) and in Active Masturbation it must surely be my partner who ejaculates. But what if Mutual Masturbation where if ejaculation occurs it could be alternatives 2, 3 or 4? Or consider a situation where my partner is the active partner in anal intercourse with me, ejaculates in me and produces ejaculation in me without any manual stimulation on my part? These considerations made it essential to construe the ejaculation as a two-place event.

Condoms and Other Destinations

Similar sorts of problems arose in trying to conceptualize the use of condoms, (especially in anal intercourse) and destinations of the sperm other than in the partner's mouth or anus. In the usual descriptive context, this may not be important, but thinking through the different consequences for HIV transmission made us realize that there must be no ambiguity about what happened to the sperm after ejaculation. After various trials, it was decided to modify the ejaculation to allow five, instead of two, possible outcomes: into a person, onto a person, into a condom, some other destination and no ejaculation.

Fuller Form of Ejaculation Codes:

As in the above short form, there is a two-place code, for EGO and ALTER, but instead of just coding whether or not ejaculation occurs, the detail of the "destination" of the ejaculate is also described for both partners, using the codes:

17This fuller form (referred to elsewhere [Coxon, 1988a] as "Chriscode" after its originator Chris Mitchell) is the form currently used to encode Diaries in project SIGMA.
The code is used to represent (in the first position) the destination of Ego’s ejaculation and (in the second position) to Alter’s ejaculation, and follows the behaviour and modality specification, separated by a comma, i.e.:

**Modality Behaviour, Ego’s Ejaculate, Alter’s Ejaculate**

**Active “Wank”**

i.e., I masturbated him with my ejaculation on him and he did not ejaculate. Other examples of its use might include:

<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>PENAM</td>
<td>Passive (“fucking”): Ego No ejaculation, Alter ejaculation in Me, with lubricant used.</td>
</tr>
<tr>
<td>IF&amp;SWXC</td>
<td>Simultaneous passive (“fucking”) and solo (“wanking”), with Ego ejaculating “elsewhere” (not in or on the partner) and Alter ejaculating in condom in Ego</td>
</tr>
<tr>
<td>SN</td>
<td>Solo (“wanking”), no orgasm</td>
</tr>
</tbody>
</table>

This extended code is normally too complex for use by subjects, and only the most resolute attempt it. In the current Notes, the simpler form is used, but all computer coding is done with it.

Before putting all these components together to lay out the structure of sexual behaviour, three further issues must be briefly discussed: the representation of *complex sexual acts*, the representation of their sequencing, and the description of the *accompaniments* of sexual activity.

### Complex Sexual Acts and Sequencing

On occasion, two (or sometimes more) sexual acts occur simultaneously in a way that may have implications for HIV transmission. For instance, masturbation is often combined with anal intercourse, but in very different forms. In this case, we introduce *concatenation* as an operator to combine two acts done simultaneously and symbolize it by the ampersand "&". Examples commonly include:

- **AF & HW** (Ego enters Anal anus and at the same time Alter is masturbating himself).
- **PS & ATI** (Ego is being fellated and at the same time tweaking Alter’s nipples).

Complex (concatenated, simultaneous) sexual acts are treated as a single act.

### Precedents and Accompaniments

The Precedents and Accompaniments give information about the context or setting of the sexual behaviour which might have relevance for HIV transmission. Although it is useful to distinguish Precedents and Accompaniments in this way, it has a degree of artificiality since there will be some things which continue before and throughout, others which are used intermittently (e.g., nitrites or “poppers”), and yet others which will refer only to a specific act (e.g., the use of a lubricant or dildo or a condom). It is usually the more general aspects which enter into the specification of the sexual session and accompaniments can, when relevant, be encoded in the range of behaviour to which they apply. Those which are of special relevance include:

<table>
<thead>
<tr>
<th>Class</th>
<th>Examples</th>
<th>Sexual Act</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Drugs</td>
<td>Amyl butyl nitrites [I]</td>
<td>Cunnilingus, heroin, etc.</td>
</tr>
<tr>
<td>Alcohol</td>
<td>Finger stools</td>
<td>Digital-anal (Fg)</td>
</tr>
<tr>
<td>Dental dams</td>
<td>Veterinary gloves</td>
<td>Anal-rectal (R)</td>
</tr>
<tr>
<td>2. Lubricants</td>
<td>Water based</td>
<td>Ano-brachial (F)</td>
</tr>
<tr>
<td>Oil-based</td>
<td>Vegetable oil based</td>
<td>Saliva</td>
</tr>
<tr>
<td>3. Toys</td>
<td>Restrictors (&quot;cock-rings&quot;)</td>
<td>&quot;Pain-pleasure&quot; (bolts, nipple-clamps, pendant weights)</td>
</tr>
<tr>
<td>&quot;Toys&quot; [T]</td>
<td>Stimulators (dildos)</td>
<td></td>
</tr>
</tbody>
</table>

(Note: letters in square brackets (e.g., [I]) indicate Coding symbols)

Some comments about those accompaniments are in order:

**DRUGS:** The most common and extensively used drugs used in homosexual intercourse are amyl and butyl nitrites (“poppers”), used to produce a rush of blood which can enhance orgasm, and which in the U.K. can be obtained without prescription in most night-clubs. Whether they are involved as cofactors in HIV transmission is a long-protracted debate of no certain outcome. Alcohol is included since this may be used to lower inhibitions (and risk-tolerance).

**PROPHYLACTICS:** These include condoms for anal intercourse, or finger stools for digital-anal insertion or veterinary gloves for brachio-anal insertion. They are used to prevent transmission of bodily-liquids or to protect lesions. Apart from condoms, their use is relatively uncommon.

**LUBRICANTS:** Much health education has centered round distinguishing the “safe” water-based lubricants which do not weaken or destroy latex rubber (such as the proprietary KY lubricant jelly, and those which also contain spermicidal and HIV-destroying Nonoxynol 9). In the past, oil-based Vaseline and Crisco vegetable oil have been extensively used, but since they do destroy latex, their use with condoms is counted as "unsafe." It is usually easier to ask the subject to name the brand used. It is also important to include saliva among lubricants,
not only because it is used as an emergency lubricant, but also because its role as a "bodily fluid" implicates it, however weakly, as a possible agent of HIV transmission.

"TOYS": it is unlikely that most "toys" are involved in HIV transmission, but some may produce small lesions, either externally (nipple-clamps) or internally. (Since many objects can be used for anal and vaginal insertion, it is worth asking which particular objects have been used.)

Putting the Structure Together

Having now described the components of the Sexual Session and Sexual Act, we need to specify the inter-relationships and structure of the components before explaining their applications in the form of the Sexual Diary and the Inventory of Sexual Behaviour.

The formal specification (syntax) of the structure has been described in Backus notation in Appendix 2, and this also serves to define economically the coding conventions explained below.

In brief, the basic elements are (open, extendible) lists of sexual behaviours, types of modality ("modes") and outcomes ("orgasm/ ejaculation"), together with accompaniments ("object" as modifiers). From these are then built the main building blocks of sexual activity: the sexual acts, which in turn are part of a sexual (sub)session.

The utility of this formalism can best be explained in the context of the Coding scheme and is therefore deferred to the following section.

Encoding Sexual Activity

The sexual Coding Scheme evolved by Project SIGMA can be used in several ways: as a useful shorthand to describe actual sequences of sexual activity (observed directly or on a video, etc); as a recondite or cryptic reference to sexual activity; as a simple and unambiguous encoding for data entry; as a way of specifying sexual Inventories and interview questions; and for comparing different versions and accounts.

The Coding scheme is designed not only for sexologists but also for subjects and respondents themselves, who (perhaps surprisingly) find it relatively simple to use. Some subjects will prefer to adopt part of the code and use shorthand description for more complicated acts, and this is a perfectly acceptable use. Indeed, we found that codes for more common activities passed into the everyday argot in some communities where we were researching.

In explaining the code, it is best to concentrate first on defining sexual acts. These consist basically of M x B x O (Modality types by Behaviours by Outcome types). In principle, this means that there are 5 x 4 acts recognized as distinct, depending on the number of behaviours chosen for inclusion. In practice, there are less, since not all modalities apply to each act (either as logically impossible [e.g., SWX = SW,NX] or physically impossible [e.g., MFJ, mutual "fisting"], or if one (e.g., SP, anal intercourse with oneself) is irrelevant). (But it is important not to decide a priori that something is impossible; ingenuity and physical flexibility and endowment can produce surprising variants; self-fellatio with ejaculation into oneself [SSO = SSM] and ejaculation into one's own mouth as a result of masturbation [SMW] have occurred in our sexual diaries.)

The sexual act is normally made up of one modality, one behaviour and one outcome (in that order). If there is no outcome, it may be omitted. Examples of simple sexual acts are:

<table>
<thead>
<tr>
<th>SEXUAL ACT:</th>
<th>Code</th>
<th>Short</th>
<th>Extended</th>
<th>Direct</th>
<th>Translation</th>
<th>In Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: SIMPLE ACTS:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. SW (= SWN)</td>
<td>Self-&quot;wank&quot;</td>
<td></td>
<td></td>
<td></td>
<td>I masturbated myself (no ejaculation)</td>
<td></td>
</tr>
<tr>
<td>2. SWO (= SWX)</td>
<td>Self-&quot;wank&quot;/me/organ</td>
<td></td>
<td></td>
<td></td>
<td>I masturbated myself to ejaculation</td>
<td></td>
</tr>
<tr>
<td>3. PF (= PFMN)</td>
<td>Passive-&quot;fuck&quot;</td>
<td></td>
<td></td>
<td></td>
<td>I was awfully entered (neither ejaculated)</td>
<td></td>
</tr>
<tr>
<td>4. PFO (= PFNM)</td>
<td>Passive-&quot;fuck&quot;/him/organ</td>
<td></td>
<td></td>
<td></td>
<td>I was awfully entered and my partner ejaculated in me</td>
<td></td>
</tr>
<tr>
<td>5. PFNM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>I was awfully entered and I ejaculated as a result of this stimulation</td>
<td></td>
</tr>
<tr>
<td>6. AR</td>
<td>Active-&quot;rim&quot;</td>
<td></td>
<td></td>
<td></td>
<td>We simultaneously fellated each other (&quot;69&quot;) and both came to orgasm</td>
<td></td>
</tr>
<tr>
<td>7. MSOC MSHM</td>
<td>Mutual-&quot;suck&quot; me in him &amp; he in me (both orgasms)</td>
<td></td>
<td></td>
<td></td>
<td>We simultaneously fellated each other (&quot;69&quot;) and both came to orgasm</td>
<td></td>
</tr>
</tbody>
</table>

| B: COMPLEX ACTS: | | | | | | |
| 8. AF&HW | Active-"fuck"-AND-He-"wank" | | | | At the same time: I orally entered him and he masturbated himself (neither to orgasm) |
| 9. PS&AHW | Passive-"suck"-AND-Active-"wank", me to orgasm in him, him not | | | | At the same time: I was fellated (came to orgasm and masturbated him (not to orgasm) |
| 10. ARHAW | Active-"rim"-AND-Active-"wank" | | | | At the same time: I orally stimulated his anus and masturbated him (without orgasm) |

1 The use of a shorthand code for sexual activity is probably best known in the context of the Kinsey studies (Pomeroy et al., 1965), but others exist (e.g., Marks et al., 1986). The history of erotic cryptography is fascinating in its own right (McCormick, 1980).

2 This is far from a trivial use. When respondents keep sexual diaries in full form, they may refer to illegal activity or be viewed as unacceptable or pornographic by others, and the consequences may be dire, as in the case of a subject who was summarily ejected from his apartment when his landlord discovered the (longhand, explicit) sexual diary he was keeping for the project.

3 A good example is SWO (Self-Wank to Orgasm), which was so widely known in South Wales in the gay community that we used the slogan "Why SWO when you can be interviewed by Project SIGMA?" as a recruiting slogan and as a T-shirt design.

4 Because SWXO would assert that I masturbated myself (SW) and, as a direct result of this, my partner ejaculated (XO). On the other hand, SW & HWXO would correctly encode a "parallel" masturbation where only my partner came to orgasm.

5 The code is given in the simpler form and the extended form second, so in this case SWXO means "solo-wank, I didn't come he did come to orgasm" and SW,NX means "solo-wank, no my ejaculation his ejaculation stimulation other" which are identical.
Comments on these examples will help in explaining some of the more subtle points of the coding scheme:

Examples 1 & 2 illustrate the fact that it is not necessary to specify the "outcome" code when ejaculation did not occur. If necessary, it can be specified as a "null outcome," as in SW, N (or, indeed, SW, N, N) as equivalent to SW.

Examples 3, 4, 5 deal with different variants of receptive and intercourse. #3 is unproblematic. In #4, the shorter form is not ambiguous, unless it was an attempt to encode the fact that it was I that came to orgasm, not my partner. The fuller form makes the original intention clear and #5 encodes a different variant: that whilst my partner did not ejaculate as a result of anally entering me, I did. Another alternative—that I ejaculated as a result of masturbating myself whilst being anally entered—can be coded by: PPFSWXX (= PP & SW, N, N).

Example 6 is unproblematic. Note that it would be possible (but unusual) for oral-anal contact to result in the recipient's orgasm: ARLNX.

Example 7 illustrates an instance where shorter form (i.e., MSO) would be ambiguous, since it would not be clear which partner came to orgasm as a result of the mutual fellatio. The fuller form makes it clear that both do in this instance.

Example 8 is a not uncommon happening (like its complement PF & SW). No ejaculation occurs in this case, but if it does, the details can be exactly coded using the extended code, and if necessary the details of the coital position adopted can be encoded by using a supplementary code.

Example 9 illustrates an instance of ejaculation occurring within a complex act.

Example 10 provides a case of a person doing two sexual acts at the same time—in this case both orally stimulating the partner's anus and masturb员ating him at the same time.

The Sexual Diary

As will by now be apparent, the structure and coding scheme was originally developed in the context of collecting sexual diaries from sample members of Project SIGMA, 24 where it was important to ensure that the content (highly sexually explicit and sometimes referring to possibly illegal activity) was incomprehensible to outsiders; since the contents could also be construed as "pornography" returned through the U.K. postal system—also possibly illegal—it was important to make the contents as opaque as possible, whilst still being comprehensible to the subject.

During the course of each SIGMA interview, the subject was taken retrospectively through the previous seven days (starting at that day) and asked about his sexual activity. These data were elicited in the framework of the structure presented in this paper. 25 In this way, the subject was taught what aspects of his sexual behaviour were relevant for the diary, and a week's data were obtained. He was then asked to keep a Sexual Diary on a daily basis for the coming month and if he agreed was given a set of Notes on Keeping a Sexual Diary, a set of diary forms and Postcard return envelopes. In the instructions, it was stressed that even partially filled diaries were useful, and respondents were encouraged to write the diary either in longhand or in code, or in a mixture of both, so long as the essential elements were included.

Sexual diaries were also solicited from the readers of Gay Times and other publications.

An example of how a longhand session account can be encoded now follows:

LONGHAND REPORT

We'd been drinking in a gay bar in Amsterdam, where we met. He was 30s, from Utrecht, into leather. We went back to my hotel room just after 1 a.m. and after sharing a joint, went to bed... (DIARY ENTRY)**

MDK AND FS MSXO (AFX) 11 HWXX 1

MDK

Obviously, the subject would not have encoded his own diary in this manner, and only a cognoscento could do so. There is usually, therefore, an editing phase to check a subject's coding or encode a longhand version. Details on editing and software representation are contained in a later section together with discussions on reliability and validity of this method.

** (Alternative extended coding would be: MDK AS PS MS, NM (AF, CN, HW, NW) 1)

24 Project SIGMA Coital Position Coding scheme uses iconic symbols to encode sexual positions, especially for anal (and vaginal) intercourse. For example, a symbol like "#" is used to encode partner #1 as inserter and standing up, with #2 the receptive partner, bent at the waist.

25 Fuller details are contained in [Davies and Coxon, 1990], q.v. Additional document.
The Inventory of Sexual Behaviour

The Inventory of Sexual Behaviour (ISB) is a method of collecting systematic information about sexual behaviour which implements the structure of sexual behaviour presented in this paper. It is not so much a single Research Instrument as a method for constructing any number of such Inventories which have the important quality of being mutually compatible and comparable. The chief advantages of the Inventory are that it is flexible, adaptable to different methods of data collection and settings; it can easily be adapted to differing lengths and complexity and to new behaviours; it is gender and orientation-free, usable for heterosexual, bisexual, and homosexual populations; and finally, it is adaptable to different time periods.

FLEXIBILITY

The ISB was designed primarily for use in an interview situation but was quickly adapted to use both as a self-administered instrument and as a Clinically based checklist for use in taking a sexual history. It is now used routinely not only in Project SIGMA and associated projects and Clinics, but is also used by all national studies as part of the Core (Mandatory) section of the Question Schedule of the WHO/IPA Homosexual Response Studies. Because it is based on the simple structure of the SBC (Sexual Behaviour Code), it is easily memorable and can be (and is) used to report sexual behaviour in covert contexts such as locations of casual sexual encounters.

Forms of the ISB vary from the simplest (designed to elicit information on anal intercourse only, which is 4 items long), through the basic heterosexual homosexual behaviour (covering the four most common behaviours; 16 items long) and the fullest (which explores the 3 principle modalities for 6 main behaviours, looking at different genders of partners, condom use and differential behaviour between partners; this amounts to 720 items, but averages a completion time of only 18 minutes).

GENDER AND ORIENTATION FREE

As far as possible, the items are kept as free of gender or orientation-marking as possible. This is most often done by including the sex in the structure of the item (“... with a man?”, “with a woman?”). Obviously, a sexual behaviour such as vaginal intercourse is intrinsically heterosexual marked, but variants such as fellatio and cunnilingus can be dealt with as “oral-genital” sex. Some problems also arise in terms of outcomes, since only males ejaculate, but again, this can be generalised to "orgasm" if desired.

ADAPTABLE

Two of the most frustrating aspects in comparing sexual behaviour studies concern the differing time periods over which subjects are asked to make an estimate (e.g., in the last week, “in the last month,” “in the last year,” etc.) and whether the response options provided are categorical (“Ever/Never”)26 or numerical (“How many times ...?”). In the ISB, both these factors can be varied according to wish.

Question Format in the ISB

The simplest question format involves Behaviour and Modality only, e.g.,
"Have you ever masturbated?"27 yourself?

but (probably) the most complex structure would be illustrated by:
"How often in the last month have you (masturbated/year [male] partner number" "in this ejaculation" "with him using a condom?"

which involves seven factors: Frequency, Behaviour, Modality, Outcome, Gender, Partner Number and Condom (though not in that order). We have seen that Behaviour, Modality and Outcome make up a Sexual Act, so the only additional elements are Frequency, Gender and Partner number.

Any variant of the ISB can be described by the combination of which of these factors are used, and what elements are chosen for each factor:

<table>
<thead>
<tr>
<th>Factor</th>
<th>Name</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. F</td>
<td>Frequency</td>
<td>Essential. Alternatives: (Categorical (Ever) or Numerical (How often) + Time period)</td>
</tr>
<tr>
<td>2. B</td>
<td>Behaviour</td>
<td>Essential. Number of behaviours at choice</td>
</tr>
<tr>
<td>3. M</td>
<td>Modality</td>
<td>Active and Passive essential. Usually including Mutual, and sometimes Self</td>
</tr>
<tr>
<td>4. O</td>
<td>Outcome</td>
<td>Optional. Ejaculation or orgasm</td>
</tr>
<tr>
<td>5. G</td>
<td>Gender</td>
<td>Optional (though to omit this factor presupposes a given sexual orientation)</td>
</tr>
<tr>
<td>6. C</td>
<td>Condom</td>
<td>Optional. Allows the distribution of sexual acts over different partners to be investigated, and can include &quot;casual&quot; as a generic alternative.</td>
</tr>
<tr>
<td>7. P</td>
<td>Partner number</td>
<td></td>
</tr>
</tbody>
</table>

Using these factors, it is possible to describe concisely any variant of the ISB by specifying what factors and constituent components go to make up the particular version. In order to make the specification exact, it is sometimes useful to add an INCLUSION or EXCLUSION list detailing any peculiarities of a given version. Given the specification, it is simple to generate the Inventory itself. Thus, the simplest (male) ISB (detailed in Appendix 4) can be defined as:

F x B x M x G x O, with
F = (Ever; Last month (0))
B = (Masturbation, Fellatio, Anal Intercourse)
M = (Active, Passive)
G = (Male, Female)
O = (With, without orgasm)
and the inclusion list:
INCLUDE (SW, SOW, AVF, AVFO).

26 An interesting variant which we have used in SIGMA is the trichotomy: (Ever/Once only/sometimes) since for some sexual behaviours there is the significant category of “I tried once, only once, and I didn’t like it, so never again.”

27 A reminder: terms in curly brackets (I am subtstituted by the subject’s preferred term when administered verbally.)
Given that specification, the question format can be generated as:

HAVE YOU:
EVER ON THE LAST MONTH
(MASTURBATED)
YOURSELF
TO YOUR ORGASM?
which with "street" substitution might read as:
"Have you ever wanked yourself? ... and come?" A more complex version also exists (and is available upon request). This is the shorter version of the World Health Organization Homosexual Response Studies Core Scale, defined as:

F x B x M x O x G x C, with
F = (Ever, last six months)
B = ( Masturbation, Fellatio, Anal Intercourse, Vaginal Intercourse)
M = (Active, Passive)
O = (With, without Condom)
G = (With, without Ejaculation)
C = (With, without Condom).

(Inclusion (SW, SWO; ARI [FM], PRI [FM]).

(Note that the WHO ISB also includes in O the variant "where he pulled out before ejaculation," i.e., the "On" of the longer Outcome list).

Issues in the Reliability and Validity of Sexual Diaries and the Inventory

At this stage, the test/re-test reliability of the Inventory is still under investigation, and the overall correlation averages 0.84 for the Shorter version (and 0.93 for rank-order correlation). Full reliability tests will be made for the various components/behaviours (because reliability is higher, for instance, with the most salient act of anal intercourse than for masturbation) and for different length variants. (For example, for those with more than one regular partner, reliability is higher within each partner relation.)

Reliability of the Sexual Diaries is rather more difficult to test, since no subject can be expected to keep two diaries in parallel and without comparison. However, the one-week retrospective Interview Diaries provide a more reasonable proposition, and in the current wave of 1983 SIGMA interviews, a random subset of subjects is being asked to repeat the information later in the interview. Nonetheless, cross-method validation is not so difficult since subjects are also asked to estimate the frequency with which they have performed the various acts over month-periods, and in some instances this month and the Diary month are identical. Earlier pretests (reported in Coxon, 1988a) indicated that Diary/Inventory estimates are generally in quite good monotonic agreement, but subjects differed markedly but consistently in under- or over-estimation, and this carries over into the estimation for other factors such as the number of sexual partners.

A more interesting variant of checking the validity of the Sexual Diary data occurs when both members of an exclusive relationship are project subjects and are interviewed either at the same time or sequentially, so that no collaboration or collusion is possible (see Abramson, 1988, p. 338). At present, our investigations of this are not systematic, but preliminary results are interesting, suggestive and encouraging.

An example is the following case. The subjects are a sexually exclusive couple in their late twenties who co-habit and had been in the relationship six months at the time of the data collection. They were interviewed in their house at the same time by separate interviewers, and the interview included the retrospective week-long Sexual Diary, taken by the interviewer. The couple's sexual week is reproduced (in the interviewers' simpler version of Code) in parallel below:

<table>
<thead>
<tr>
<th>Day</th>
<th>Subject #1</th>
<th>Subject #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>WED</td>
<td>pm, P1: MDK AFOX</td>
<td>pm, P1: MDK PFXO</td>
</tr>
<tr>
<td>THURS</td>
<td>pm, P1: MDK AFO</td>
<td>(Cannot remember)</td>
</tr>
<tr>
<td>FRI</td>
<td>pm [5 in L*] P1: MDK AFOX</td>
<td>(Cannot remember)</td>
</tr>
<tr>
<td>SAT</td>
<td>am P1: AF (then shopping)</td>
<td>am P1: PF</td>
</tr>
<tr>
<td>SUN</td>
<td>pm P1: AFOX</td>
<td>pm P1: PFXO</td>
</tr>
<tr>
<td>MON</td>
<td>am P1: MDK MWOO</td>
<td>am P1: MDK MWOO</td>
</tr>
<tr>
<td>TUES</td>
<td>pm P1: MDK MWOO MDk Allo PdS MWs PW APl</td>
<td>pm P1: MDK MWOO</td>
</tr>
<tr>
<td>TUES</td>
<td>PM AW AE MWOO</td>
<td>Pri PFOO</td>
</tr>
<tr>
<td>TUES</td>
<td>pm SWOP</td>
<td>pm MDK MWOO</td>
</tr>
</tbody>
</table>

(IS in L*) means that the subject was dressed in leather.

Several points here are worth comment. All Subject #2’s reported acts are compatible with those of his partner, as are those in Subject #1’s slightly fuller account, which includes an additional three (non-solitary) acts, and two days which Subject #2 could not recall. In this example, anal intercourse is entirely single-role: if Subject #1 engages in anal intercourse he is always inserter, if Subject #2 does, he is always inserter. Masturbation, on the other hand, is varied—either mutual, active or passive—and fellation does not occur in this week. Their repertoire also includes some “light S&M” activities: “Water-sports” (Urination on), “Bondage” (Tying up), and the use of leather. In terms of patterns (Davies 1989), the sex is therefore mostly role-ist (dominant/active vs dominated/pasive) rather than reciprocal, and the sexual sessions are all end-marked by ejaculation (more often on Subject #1’s part).

It should be repeated that agreement is not usually as good as this, and one often encounters sessions with misplaced days and different detail. But there does appear to be considerable structural similarity in accounts (role/reciprocal and ejaculation end-marking).

Representation and Analysis

The representation and analysis of the ISB are straightforward matters. They produce rectangular data (N subjects by p variables), are simply represented in that format by any conventional data analysis package such as SPSS or SAS, and may be analyzed by any relevant statistical or test-analysis procedures. The only caution is that since such data are often only ordinaly reliable, caution should be exercised in their analysis by parametric and linear
models, with a preference being for non-metric procedures (Tukey data exploratory analysis and ordinarily invariant methods such as multidimensional scaling (see Coxon, 1982)).

**Sexual Diaries**

Diary information, on the other hand, yields challenging and novel data. Although special-purpose software packages are now available, and preliminary analysis has usually to be done on the diary data before they become susceptible to further analysis. The strategy followed in Project SIGMA is to employ a simple, easy to use, “flat” database (CARDBOX-PLUS) as a data input medium and use it for simple logical checks and analyses and counts, and then output the results in a machine-readable form for later analysis by purpose-written program packages.

In the database context, the sexual session becomes the RECORD, and the sexual act(s) the main field (ACT) with additional fields for identifiers, partner references, general-purpose retrieval variables (age, site, relationship-type) and important tags (e.g., whether or not condoms were used). As an illustrative example, the sexual session from the previously described “Longhand Report” is encoded in cardbox as follows:

<table>
<thead>
<tr>
<th>NO CF/87</th>
<th>TYPE II</th>
<th>STATUS: N</th>
<th>DAY: Fri</th>
<th>DATE: 26/06/86</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIME: 12.00</td>
<td>1PLACE: My hotel room, Amsterdam</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERSON: CS (30s, male, Utrecht, Into leather)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT: AS MK AS MS, NM (AF/CN/HW/NH/PH MDK)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEPERS: Y</td>
<td>1CONDOMS: Y</td>
<td>ILUS: Y (ky)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER: Met in gay bar; smoking marijuana</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The main shortcoming of a “flat” data base is that it is non-relational, and therefore there cannot be direct linkage to a Partner file. However, such a record can later be exported to a database file and given partner relation. However, much simple exploratory analysis of such data consists of finding examples of particular behaviors and patterns of behaviour, and then counting them (since the record is normally the session rather than the individual, the unit of counting is the session; sessions have to be aggregated up to the individual to get individual incidence data). An example of the logical instructions in CARDBOX to obtain the records that contain an instance of “receptive anal intercourse where a condom was not used” is given below.29

---

29In CARDBOX instructions, fields (such as Acts and Condoms) are selected and matched by a frame. The symbol “*” means “any sequence of letters.”

---

<table>
<thead>
<tr>
<th>CARDBOX-PLUS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>File contains 844 records</td>
<td></td>
</tr>
<tr>
<td>Level 1: SELECT AC4F* 80 records selected</td>
<td>80 ACs contain a behaviour with F (i.e., including anal intercourse)</td>
</tr>
<tr>
<td>Level 2: SELECT CO/NN 40 records selected</td>
<td>of these, 40 did Not use a Condom</td>
</tr>
<tr>
<td>Level 3: EXCLUDE AC4AF* 20 records removed</td>
<td>removing 20 instances of active “fucking,” leaves only passive</td>
</tr>
</tbody>
</table>

I.e., of the 844 sexual sessions, 20 (2.4%) were instances of receptive (passive) anal intercourse where a condom was not used.

The linguistic analogy for the structure of sexual behaviour is more than that. Because Sexual Sessions are “sentences” and Sexual Acts are “words,” and these words are prefixed and inflected, these data can be analyzed by programs designed for linguistic and content analysis. In particular, facilities such as KWIC (“Key Word in Context”) are excellent for examining the context and immediate neighbourhood of chosen activities. Taking receptive anal intercourse again, the following (part) of a KWIC analysis of some of our 1986 Diary data make this clear:

<table>
<thead>
<tr>
<th>ID</th>
<th>HIV</th>
<th>Date</th>
<th>SEXUAL SESSION</th>
<th>Partner #</th>
</tr>
</thead>
<tbody>
<tr>
<td>027</td>
<td>NT</td>
<td>11/10/86</td>
<td>AS PS MW PF AFs</td>
<td>P2</td>
</tr>
<tr>
<td>0016</td>
<td>AF</td>
<td>12/10/86</td>
<td>AW PS PF PP</td>
<td>P2</td>
</tr>
<tr>
<td>0017</td>
<td>AF</td>
<td>14/10/86</td>
<td>AFPP PEO AF PS</td>
<td>P1</td>
</tr>
<tr>
<td>0018</td>
<td>AF</td>
<td>01/06/86</td>
<td>AS AF PS</td>
<td>X5</td>
</tr>
<tr>
<td>0019</td>
<td>AF</td>
<td>05/06/86</td>
<td>PEO PS</td>
<td>X4</td>
</tr>
<tr>
<td>0020</td>
<td>AF</td>
<td>22/06/86</td>
<td>AS PEO</td>
<td>P1</td>
</tr>
<tr>
<td>0021</td>
<td>AF</td>
<td>14/06/86</td>
<td>AFO PEO</td>
<td>X5</td>
</tr>
<tr>
<td>0022</td>
<td>AF</td>
<td>20/06/86</td>
<td>PEO AF</td>
<td>P1</td>
</tr>
<tr>
<td>0023</td>
<td>AF</td>
<td>24/06/86</td>
<td>PEO AF</td>
<td>X1</td>
</tr>
<tr>
<td>0024</td>
<td>AF</td>
<td>01/06/86</td>
<td>AFO PEO</td>
<td>P1</td>
</tr>
<tr>
<td>0025</td>
<td>AF</td>
<td>07/06/86</td>
<td>AFO AF</td>
<td>P1</td>
</tr>
<tr>
<td>0026</td>
<td>AF</td>
<td>18/06/86</td>
<td>PEO AF</td>
<td>P1</td>
</tr>
<tr>
<td>0027</td>
<td>AF</td>
<td>23/06/86</td>
<td>PEO AF</td>
<td>P1</td>
</tr>
<tr>
<td>0028</td>
<td>AF</td>
<td>30/06/86</td>
<td>PEO AF</td>
<td>P1</td>
</tr>
<tr>
<td>0029</td>
<td>AF</td>
<td>07/06/86</td>
<td>PEO AF</td>
<td>P1</td>
</tr>
<tr>
<td>0030</td>
<td>AF</td>
<td>13/06/86</td>
<td>PEO AF</td>
<td>P1</td>
</tr>
<tr>
<td>0031</td>
<td>AF</td>
<td>18/06/86</td>
<td>PEO AF</td>
<td>P1</td>
</tr>
<tr>
<td>0032</td>
<td>AF</td>
<td>08/12/86</td>
<td>PEO AF</td>
<td>P1</td>
</tr>
</tbody>
</table>

By the use of KWIC, a person’s (or a group’s) behaviour can be examined in terms of background factors such as the distribution by HIV status and within person in terms of his different relationships (“F” here means regular partner and “X” a casual partner).

Patterns are also easier to discern in this format; without over-interpreting these sample (but genuine) data, it looks as if:

- 027 tend to longer or more varied sessions when receptive anal intercourse is involved and that being falled to his ejaculation (PWO) is an end-marker for him, whereas
- for X025 (who is HIV positive) PFO (or PFOO where both come to ejaculation), is the end-marker, whoever the partner is, and sometimes it is the only activity.
CONCLUSION

In this paper, we have provided a specification of the structure of sexual activity sufficient to represent detailed sexual behaviour sequences which are relevant to the transmission of HIV. The structure is both close enough to subjects' accounts to enable simple use and it is precise enough to allow it to be represented formally as a language-like simple algebra. Finally, the structure thus defined provides the means for defining research instruments of very different forms (inter alia systematic inventories, Life Histories and Diaries) which are both flexible and comparable. We therefore commend them to the research community.

References


Details of the version 2.0 of the Project Sigma SEXDIAIR program are available from the author.


**APPENDIX ONE**

In Question Schedule 1 of Project SIGMA, subjects were asked to give their "preferred name" for a range of common sexual terms. These were then substituted in questions using the terms. Although there was some differences in usage by area (e.g. South Wales had unique terms like "booking" for (anal) intercourse), distributions were markedly similar.

### Preferred Terms for Sexual Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penis</td>
<td>53%</td>
</tr>
<tr>
<td>Cock</td>
<td>18%</td>
</tr>
<tr>
<td>Prick</td>
<td>15%</td>
</tr>
<tr>
<td>Penis</td>
<td>8%</td>
</tr>
<tr>
<td>Dick</td>
<td>5%</td>
</tr>
<tr>
<td>Willy</td>
<td>1%</td>
</tr>
<tr>
<td>Other</td>
<td>1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semen</td>
<td>44%</td>
</tr>
<tr>
<td>Come/Cum</td>
<td>30%</td>
</tr>
<tr>
<td>Spunk</td>
<td>8%</td>
</tr>
<tr>
<td>Sperm</td>
<td>5%</td>
</tr>
<tr>
<td>Semen</td>
<td>4%</td>
</tr>
</tbody>
</table>

### B. Preferred Terms for Sexual Behaviours

<table>
<thead>
<tr>
<th>Term</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masturbation</td>
<td>77%</td>
</tr>
<tr>
<td>Pull/Toss Off</td>
<td>4%</td>
</tr>
<tr>
<td>Other</td>
<td>1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fellatio (Oral Sex)</td>
<td>61%</td>
</tr>
<tr>
<td>Blow-Job</td>
<td>21%</td>
</tr>
<tr>
<td>Fellatio</td>
<td>8%</td>
</tr>
<tr>
<td>Oral Sex</td>
<td>2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anal Intercourse</td>
<td>43%</td>
</tr>
<tr>
<td>Screw</td>
<td>20%</td>
</tr>
<tr>
<td>Intercourse</td>
<td>10%</td>
</tr>
<tr>
<td>Other</td>
<td>4%</td>
</tr>
</tbody>
</table>

---

**APPENDIX TWO: Syntax of Sexual Structure Codes**

- `<SESSION>` ::= {
  `<ACT>` | `<SUBSESSION>` | `<CONJUGATOR>` | `<SESSION>` }
- `<SUBSESSION>` ::= {
  `<ACT>` | `<SESSION>` | `<MODIFIER>` |
- `<ACT>` ::= {
  `<MODE>` | `<BEHAVIOUR>` | `<EGO's ORGASM>` | `<ALTER's ORGASM>` | `<MODIFIER>` |
- `<MODE>` ::= {
  `SHTAIPIM` |
- `<BEHAVIOUR>` ::= {
  `WISIFIR11Fg1Dk1F1 ...` |
  `<EGO's = ALTER's ORGASM>` ::= {
  `OIIIEICIK` |
- `<MODIFIER>` ::= {
  `null / <associated object list> <modifier>` |
  `<OBJECT>` ::= {
  `P.L.D.T. ...` |
  `<CONJUGATOR>` ::= {
  `space + 1` |

**Notes**

1. The symbol "::=
   may be read as "can be replaced by" or "consists of": it links the basic term (definendum) on its left hand side and its specification on the right. The symbol " := " may be read as "is" or "such that". The most fundamental units (behaviours, modes, orgasm/jaculations, conjugators and objects or "accompaniments") are specified as a list of letters denoting the contents in the code.

2. The more complex units (session, act, modifier) are built out of these units, so that the hierarchy is:

   `{behaviours, modes, outcome/s} => {acts} => {modifiers, objects}
   (conjugator)
   (sub)SESSIONS`.

3. I am grateful to Chris Mitchell with whom we developed the first version of this specification.
<table>
<thead>
<tr>
<th>Questions and filters</th>
<th>Coding categories</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sexual Behaviour</strong></td>
<td></td>
</tr>
<tr>
<td>In the past six months have you ever done the following sexual acts?</td>
<td></td>
</tr>
<tr>
<td><strong>INVENTORY OF SEXUAL ACTS</strong> (SHORT FORM)</td>
<td></td>
</tr>
<tr>
<td>Have you:</td>
<td></td>
</tr>
<tr>
<td>(A) [MASTURBATED]</td>
<td></td>
</tr>
<tr>
<td>yourself</td>
<td></td>
</tr>
<tr>
<td>without (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>to (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>a female?</td>
<td></td>
</tr>
<tr>
<td>a male</td>
<td></td>
</tr>
<tr>
<td>without (his) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>to (his) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>BEEN [MASTURBATED] BY</td>
<td></td>
</tr>
<tr>
<td>a female</td>
<td></td>
</tr>
<tr>
<td>without (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>to (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>a male</td>
<td></td>
</tr>
<tr>
<td>without (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>to (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td><strong>Have you:</strong></td>
<td></td>
</tr>
<tr>
<td>(B) [FELLATED]</td>
<td></td>
</tr>
<tr>
<td>a female? (cunnilingus)</td>
<td></td>
</tr>
<tr>
<td>a male</td>
<td></td>
</tr>
<tr>
<td>without (his) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>to (his) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>where he pulled out before [ejaculating]</td>
<td></td>
</tr>
<tr>
<td>where (he) wore a condom?</td>
<td></td>
</tr>
<tr>
<td>BEEN [FELLATED] BY</td>
<td></td>
</tr>
<tr>
<td>a female</td>
<td></td>
</tr>
<tr>
<td>without (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>to (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>a male</td>
<td></td>
</tr>
<tr>
<td>without (your) [ejaculation]</td>
<td></td>
</tr>
<tr>
<td>to (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td>where you pulled out before [ejaculating]?</td>
<td></td>
</tr>
<tr>
<td>where (you) wore a condom?</td>
<td></td>
</tr>
</tbody>
</table>

<p>| <strong>SEXUAL BEHAVIOUR STUDY: CORE QUESTIONS SCHEDULE</strong>                                  |                   |</p>
<table>
<thead>
<tr>
<th>No.</th>
<th>Questions and filters</th>
<th>Coding categories</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(C) [ANALLY ENTERED]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a female</td>
<td></td>
</tr>
<tr>
<td></td>
<td>without (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to (your) [ejaculation]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(you) with a condom?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(you) without a condom?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>where (you) pulled out before [ejaculating]?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a male</td>
<td></td>
</tr>
<tr>
<td></td>
<td>without (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to (your) [ejaculation]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(you) with a condom?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(you) without a condom?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>where (you) pulled out before [ejaculating]?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BEEN [ANALLY ENTERED BY]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a male</td>
<td></td>
</tr>
<tr>
<td></td>
<td>without (his) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to (his) [ejaculation]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(him) with a condom?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(him) without a condom?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>where (he) pulled out before [ejaculating]?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Have you:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(D) [VAGINALLY ENTERED]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a female</td>
<td></td>
</tr>
<tr>
<td></td>
<td>without (your) [ejaculation]?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to (your) [ejaculation]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(you) with a condom?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(you) without a condom?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>where (you) pulled out before [ejaculating]?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Have you:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(E) [CORAL-ANALLY STIMULATED]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a female</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a male</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BEEN [CORAL-ANALLY STIMULATED] BY</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a female</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a male</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX FOUR:

INVENTORY OF SEXUAL BEHAVIOUR (male)

<table>
<thead>
<tr>
<th>HAVE YOU:</th>
<th>EVER?</th>
<th>IN LAST MONTH?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(MASTURBATED)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yourself?</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>a female?</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>a male?</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>to your orgasm?</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>to his orgasm?</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

(BEEN MASTURBATED) BY

a female? | [ ]   |
a male?   | [ ]   |

to (your) orgasm? | [ ] |

to (your) orgasm? | [ ] |

(FELLATED)

a female? | [ ]   |
a male?   | [ ]   |

to his orgasm? | [ ] |

(BEEN FELLATED) BY

a female? | [ ]   |
a male?   | [ ]   |

to (your) orgasm? | [ ] |

to (your) orgasm? | [ ] |

(ANOALLY ENTERED)

a female? | [ ]   |
a male?   | [ ]   |

to (your) orgasm? | [ ] |

to (your) orgasm? | [ ] |

(BEEN ANOALLY ENTERED) BY

a male?   | [ ]   |

to his orgasm? | [ ] |

(VAGINALLY ENTERED)

a female? | [ ]   |

to (your) orgasm? | [ ] |
Research note: Strategies in eliciting sensitive sexual information: the case of gay men

Tony Coxon
with P.M. Davies, A.J. Hunt, T.J. McManus, C.M. Rees and P. Weatherburn – Project SIGMA

Abstract

Project SIGMA's research on gay men's sexual behaviour and AIDS and HIV infection has posed complex ethnographic and methodological problems which have required unorthodox and innovative strategies for their solution. Three main problem areas included: value-conflict between interviewer/HIV tester roles; issues concerning data confidentiality and legal interference, and the elicitation of detailed sexual behaviour. The procedures developed have clear applicability to other research concerned with covert, stigmatised and legally-sensitive behaviour or data.

Project SIGMA

All involved in trying to analyze, interpret, model or simply understand the development of the AIDS pandemic have called regularly for better data relating to the modes of transmission of the HIV Virus, and in no domain more insistently that in sexual transmission (Coxon and Carballo, 1989). But few methodological innovations have been developed to help produce these better – more detailed, more reliable, more valid – data. In the West (or in the World Health Organization's terminology 'Type I' countries) quite new problems of sample definition have arisen in attempting to delineate the appropriate populations for studying high-risk behaviour. Although the appropriate populations may well be 'all male-to-male sexual behaviour' (rather than 'homosexuals') or 'intravenous substance use activities' (rather than 'drug-takers'), it is unfeasible or grossly expensive to attempt to operationalise them, and it is equally beyond the resources of conventional research.

© The Editorial Board of The Sociological Review 1993. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 238 Main Street, Cambridge, MA 02142, USA.
funding agencies to carry out population-wide studies to identify minority and socially invisible groups. Even when such general population studies are mounted, fairly trivial changes in wording or interviewer characteristics or instructions can produce prevalence estimates that differ by a factor of 2 or 3.\(^1\)

Project SIGMA\(^2\) is a set of inter-related studies with the main aim of investigating the types and patterns of sexual behaviour and HIV sero-conversion among men who have sex with men, and to look at the changes over time in behaviour, HIV status, and relevant socio-psychological variables. This involves:

- studying the men’s lifestyles using a detailed face-to-face interview at yearly intervals.
- establishing the HIV (and Hepatitis B and other viral) status using blood (or other\(^3\) body fluid) donated by the men, with appropriate pre- and post-test counselling.
- a self-completed detailed sexual behaviour diary (Coxon, 1988, 1992) filled in on a daily basis.

Together with

- subsidiary studies of casual open-air or public toilet (‘tearoom’ or ‘cottage’) sex, and of male sex-workers and sado-masochistic men.

The sampling of over 1000 men in eight sites in England and Wales poses interesting questions in its own right, but which will be ignored in this paper. Suffice it to say:

- that the two main sites were Greater London and South Wales\(^4\)
- that subjects were recruited from the naturally-occurring population and not from Genito-Urinary Medicine clinics
- that a 2-factor design was used (3 Age Groups by 3 Relationship Types - Closed, Open, No Regular) in conjunction with snowball network sampling, and
- that the longitudinal design is of (currently) five waves at yearly intervals.

### Bleeding

Originally the Project did not have taking of blood samples among its purposes; this became a requirement for the two main sites after initial discussions with the Medical Research Council, who wanted to obtain HIV-1 (then termed HTLV-3) sero-prevalence estimates among gay and bisexual men. Operating within the medical mode, the officers of the MRC assumed that this would be obligatory, and take place in a Clinic setting. But the Project had made a methodological point of not sampling from (or even, in one site, avoiding\(^5\)) the local GUM clinics, and many of the more closeted subjects would never agree to come to a Clinic, so this posed difficulties. Moreover, at this time the issue of whether to take the test was highly contentious within the gay community, and many of the most vocal and articulate gay men were strongly opposed to it. A strategy had therefore to be adopted which would maximise the number of subjects consenting to be tested and make it possible to take the blood samples in a natural (interview) setting. It was decided therefore:

- to recommend to subjects that they be tested anonymously, ie without being told the result of the test, but to arrange for the result to be given if he wished it. In the event only 13 per cent were tested anonymously.
- to train interviewers in the main sites as phlebotomists, functioning under the medical authority of the Project’s Clinical Investigator (Dr T.J. McManus) and locally under the relevant clinical authorities.\(^6\) All the Project Principal Investigators were themselves trained (and operated) as interviewers and as phlebotomists.

### Interviewers . . . and counsellors and phlebotomists

It was relatively simple to arrange for training in counselling, though it was a novelty for research staff to be told that such training and the taking of blood was part of their Conditions of Appointment.\(^7\) Arrangements had also to be made for communicating the test results, and for safeguarding the data about a subject’s sero-status.\(^8\)

The act of taking blood established a quite distinctive — and
sometimes almost proprietorial – bond between the interviewer and the subject (a good instance of Titmuss's (1973) thesis). This certainly increased rapport, but it also introduced interesting examples of role conflict. These were of various sorts.

In a few cases the conflict arose between job requirements and personal views: one interviewer disapproved of HIV testing, but agreed to take the subjects’ blood for testing as part of his job requirements. This manifestly led to a major interviewer effect both in lowering his rates of accession and of the proportion of subjects deciding to be told their HIV result.

When (as in the usual case) the same person took blood and interviewed the subject, there could be a more subtle conflict between evaluative (health) values and scientific strategy. When giving the pre-test counselling, the interviewer was obliged to articulate to the subject the ‘Safer Sex’ guidelines, under which certain activities such as unprotected anal intercourse were explicitly disapproved of as ‘high risk’. In the South Wales site it had been decided to take blood at the start of the interview, so that a pall of anxiety about giving a blood sample did not hang over it. However, an hour or so into the consequent interview, the interviewer then had to elicit information about the subject’s infractions of the very Guidelines given in the course of pre-counselling. The interviewer skills here needed to be considerable if the subject were not to ‘shade’ (or, indeed, lie about) his sexual behaviour (see Pomeroy et al. 1982). Several strategies were tried; a fairly successful one was ‘collusion’. The questionnaire [Wave 1] format is:

A lot is being talked nowadays about ‘Safe Sex’. . .

[C4.1.1] Being honest, how seriously do you take the idea of ‘safe sex’?

But in some cases interviewers prefaced C4.1.1 by words such as the following:

We know a lot of guys don’t really pay any attention to Safe Sex in the heat of the moment . . .

which of course raises questions about comparability, but illustrates the need felt by interviewers to contrast the ‘eliciting proscribed behaviour’ approach from the earlier counselling approach.10

The London Site by contrast decided to locate blood sampling at the end of the interview, and this by and large avoided the above role-conflict. But it, too, had a disadvantage. In the South Wales case, the physical reaction of the subject to the test could be monitored over the two-hour long interview, and several instances of fainting did occur during the interview as a result of the earlier blood-taking. In the London case, there was often not enough time after the end of the interview and the subject leaving to rule out the possibility of later reaction.

Confidentiality

Confidentiality of data is of course always a problem in social science research, especially when dealing with sensitive areas like sex and income. It becomes doubly sensitive when the behaviour may be proscribed and/or illegal, where the sexual orientation of the subjects is often a closely-guarded secret (whose revelation could be catastrophic). This applies to almost all studies of gay men. The problems are aggravated when the study is longitudinal (since actual identity is need for tracking purposes and subjects might reasonably have fears about the security of that information). But when, as here, the question of HIV status is added, the problems are yet further aggravated.

The first set of issues of confidentiality concern recruitment and consent: how do you convince the subject that the information he will give is safe? Any undertaking has then to be matched with project procedures which will safeguard such secrecy. The initial recruiting leaflet contained a ‘Confidentiality Guarantee’11 which promised that an individual’s identity could not be known outside the project, and when the subject came to be interviewed he was required to sign the following Statement of Informed Consent, which encapsulated a two-way contract.

STATEMENT OF INFORMED CONSENT

By my signature (or mark) below I give my informed voluntary consent to participate in a study of sexual lifestyles, conducted by: [Names and affiliations of Principal Investigators].

I understand:

— that the information will be used for research purposes only
Tony Coxon et al.

that my identity will be kept entirely confidential
that the identity of any persons I name in the interview
will be totally anonymous
that no names will be attached to any written or machine-
readable schedules.

For my part, I promise that I will give answers which are to
the best of my ability honest and accurate. I also understand
that I may withdraw my participation in the study at any time
if I choose to do so.

[Signature/mark, name address and 3rd person information]

It will be noted that the Statement promises confidentiality of
the subject's name and address; by this is meant that they appear
within internal Project records, with due security of access, and a
guarantee that it cannot be known outside. But it also promises
anonymity for the names of anyone he chooses to mention; this
means that the identity of his friends and of his partners did not
even appear on interview schedules, and in many cases was not
even mentioned to the interviewer except under labels like 'PI'
(partner 1). This was done because subjects were by and large
more concerned about protecting the identity of their friends and
nominees than protecting their own.12

Implementing the undertakings of confidentiality given to the
subjects poses more difficult issues. First, all staff signed an
Undertaking on Confidentiality which formed part of their Job
Conditions, and it was made explicit that infraction of these
undertakings could form the basis for dismissal.13

[STAFF] UNDERTAKING ON CONFIDENTIALITY

I, the undersigned, understand that in having access to the
data of the Project I agree to be bound by confidentiality in
the following ways:

1. The information on the names and identity of subjects is
entirely confidential and may neither be divulged to any
person, nor even to the subject unless he himself makes the
identification.

In particular, the (stated) HIV antibody status of a subject
must not be divulged to any person even any member of the
Project

2. Other information (on background, behaviour etc) may
never be given to any person or body outside the Project

when any individual is identifiable. In case of doubt, the
issue shall be raised with the Project Investigator.

(Signed by staff member and Investigator and dated)

(Note: Although only the Investigator knew the actual blood test
results, subjects were asked in the Interview what they knew (or
believed, or suspected) their HIV status was at the time, and
therefore interviewers were aware of the subject's (stated) HIV
status).

The same Undertaking was signed by all Interviewers and cler-
cial and data-processing staff of the Project – indeed by any and
all staff who would, or could, come into contact with data or
transcripts. With very few exceptions, the Undertaking was faith-
fully kept by staff.

Keeping track

As with any panel study, problems of attrition (and sadly of gen-
ue mortality) were considerable, and given the interest which
police and other authorities had in compiling lists of known
homosexuals (see 3.3 below), a number of subjects were unwilling
to give their actual name and address, despite the undertakings
of confidentiality. The most sensitive characteristics concern iden-
tity (name and address) and HIV sero-status. Subjects were
allowed therefore to use a pseudonym, with a warning that they
should remember what name they had chosen, as they would be
contacted in the next wave by that pseudonym. In a few cases,
subjects were not even willing to do this and in this case only the
interviewer knew their identity. At a later stage this decision was
regretted, too, when a disaffected interviewer who had left the
Project refused to give any information on such subjects and they
were lost to the panel.14

Keeping track of subjects was often easier in South Wales than
in London even though the overall response rate was lower. In
Cardiff there is only one gay club and two gay pubs, so the
'scene' is fairly compact and close-knit and by being regularly on
the scene, Project researchers often encountered their subjects
and knew how they could be reached if contact was lost. In
London, with a much more dispersed and large scene, it was
much more difficult to keep contact.

At each interview the subject was therefore asked to give the
name and address of a third person (typically a relative, partner or friend) who would know their whereabouts if they moved or lost contact. A surprisingly large number of subjects (more than 75 per cent) were prepared to do this, so long as a promise was given that any contact made did not identify the Project or its nature. Only rarely was this information actually used.

In terms of project procedures, Statements of Informed Consent (which contained Identifier, Name, Address and 3rd Person Contact information) were kept physically separate from schedules and under secure lock and key away from the project, and no data files bore anything but the code identifier.

HIV Sero-status

As mentioned in Section 2, information about the subject's HIV status came from the actual blood and saliva tests (Hunt et al. 1990) and from the subject's reported (or believed) HIV status, stated in the interview. Although HIV status information has always been sensitive, it became increasingly so in the United Kingdom as insurance and house mortgage schemes increasingly asked the 'Aids question', when even to have taken the HIV antibody test can be sufficient presumptive evidence to refuse cover or charge higher rates. Although there are signs that this policy may be moderating, it seems unlikely that a subject's open avowal of homosexual practice will be taken as anything but evidence of inherent risk. Initially, subjects often agreed to be bled only if it were anonymous; then as 'early intervention' became a favoured notion, others wanted to be bled and know their results and finally, a number agreed to be interviewed in order to get their result in a manner which would escape Insurance attention; it became increasingly important to ensure that HIV status information was secure and watertight. The manner and techniques by which this was done must remain unreported; suffice it to say that information linking identifiers and HIV status were kept on a separate machine, in a different location and fully accessible only by the Principal Investigator. The contract to respect anonymity and confidentiality brought with it a number of ethical problems for the Investigator (who knew the status and identity of those who had chosen not to know their result). For example, should a subject who is (unknown to him) HIV antibody positive and who is engaging in penetrative or other risky sex with others be forewarned? Should a subject who had been (unknowingly) positive for several years be given hints about early intervention? The answer in both cases must clearly be 'no' given the undertaking the Project gave, but in some cases it was hard to resist the temptation to hint at hidden knowledge.

The police, politics ... and paranoia?

Much of the work of the Project has been against a background of highly political issues and events. Some issues were highly specific: blood sampling for HIV antibodies was for a long time a highly contentious and divisive issue in the gay community, and views were expressed in the gay press that the Project was doing a disservice to the gay community by collaborating with scientific and medical authorities - even that by publishing clinic vs non-clinic HIV rates the Project was encouraging gay men to relax their adherence to Safer Sex Guidelines.

More potent, however, was the political and legislative background in which perceived anti-gay and homophobic reactions were embodied in 'Section 28' and 'Section 25' legislation. This formed a background when the Project was recruiting respondents or persuading them to be re-interviewed. Some were understandably concerned about their identity becoming known and/or their data accessed if they co-operated, and about the use which could be made of our scientific results by hostile authorities.

Section 28 of the Local Government Act refers explicitly to the illegality of local government authorities undertaking any support or funding of activities which could in law be construed as promoting homosexuality, or which presented non-heterosexual alliances as 'pretended family relationships'. Its chief effect was not legal action so much as a diffuse threat which blighted a wide range of activities: it was often sufficient for local councillors to threaten action for initiatives and funding to be dropped. In our case it was suggested that the research itself could be in breach of Section 28 (but at the London site only, since it is located at South Bank University which then depended on local authority funding), and it was specifically brought to the attention of the Editorial Board of the academic journal carrying our first published account of the Sexual Diary method. Section 25 refers to the Criminal Justice Bill. Among the provisions of an act which aims to 'protect the public from serious harm', victim-
less crimes relating to homosexual men are singled out for especial attention, and the 'serious harm' referred to includes the 'risk of death or serious mental or physical injury caused by further sexual offences committed by the offender', including a possible reversion to the use of compulsory psychiatric treatment for sex-related crimes. A further indication of legal background which directly affects our respondents and topics of the Project Research is the so-called 'Operation Spanner', where Police had been investigating SM videos, including some ostensibly 'snuff' videos involving bondage, corporal punishment, body-piercing and other activities. The men's defence was based on the fact that they had all consented to the activities. The judge ruled that sexual behaviour is not an area where a person can consent to an assault, despite the fact that one can thus consent to assault in games such as boxing or football, as well as in medical contexts. Moreover, if in sexual activities the pain, marks or injuries are 'more than trifling', even consensual activity is illegal. It would appear that a love-bite between a man and wife is now illegal in England and Wales, and a yet higher fraction of sexual activities of homosexual men are thus criminalised. Although the appeal judgement is likely to go to the European Court of Justice, it puts wide ranges of behaviour (which 34 per cent of our subjects say they have engaged in at some time) into the illegal category in the mean time and makes it difficult to secure truthful, explicit and detailed data on this topic.

There has been widespread fear, then, that information about identity or behaviour could be accessed or obtained from Project files - and especially by the Police authorities. Was (and is) this a justified fear?

Homosexual behaviour is illegal in Britain.20
- which is not in private, or
- which involves more than two men, or
- where either partner is under 21
- or a member of the Armed Forces (or the Merchant Navy).

On these criterion probably more than a quarter of the data of the Project could be construed as referring to illegal activity; the fear of this information falling into the wrong hands is therefore fully understandable. It was believed initially that, being under the funding andegis of the Medical Research Council, Project data had the status of 'medical records' and hence not seizable by the police. Except for the data in the possession of our Clinical Investigator, it seems unlikely that they are thus pro-
tected. But is the possibility of such access or seizure even credible? Unfortunately, yes. The Project had already encountered problems of HM Customs and Excise seizing Project training videos, and Project interview schedules being subject to confiscation by postal authorities.22

More disturbing were police attempts to obtain names and addresses of friends and contacts of a large number of gay men in their investigations of a murder in South Wales involving a gay male victim. Although this did not involve the Project directly, a subsequent investigation in Bristol alerted the police to the fact that the Project was studying gay men and might therefore have information from or about the dead man that may prove useful in their enquiries. Only because the Project was alerted to this by a phone call from the Bristol police were staff able to be fully prepared - in a manner it would be counter-productive to divulge - when asked to confirm or deny that the dead man had been an informant, and when the Police visited the Project to enquire further.23 It was not sufficient - or even sensible - to encrypt all data, since they did not include any identifier. Rather, steps were taken to ensure that the identifier/name link file could be permanently and unrecoverably erased very swiftly (ensuring that a readable, regularly updated copy disk existed in a foreign location). So far, the Project has not encountered the need to destroy data in this way, but it means that Project literature can confidently re-assure subjects about the security and confidentiality of their identity and data.

Collecting sensitive data

Questioning sexual behaviour

A large part of the Core Interview Schedule consists of detailed, systematic investigation of sexual behaviour over given periods of time, together with information on the use of prophylactics (eg condoms, surgical gloves), associated alcohol and drug use (especially nitrates, 'poppers') and 'toys' (eg dildoes). In the interview schedule, this section appears about half an hour into the interview, after preliminary sections dealing with facesheet and 'fed forward' information required for updating tracking files, and general sections concerned with identity, 'outness', partners and relationships. Data collection about sexual activity is unified by
using a single schema, the Sexual Behaviour Code (SBC), and compatible variants of this are used in many other studies (Coxon et al., 1992), including the eight national Homosexual Response Studies of the World Health Organization's Global Program on AIDS.

If such a schema is to be generally useful it must be capable of being used in a variety of data-eliciting situations. In particular, it must be applicable:

- when the subject makes his own descriptions (self-reports)
- when the sexual activity is observed (and/or participated in)
- in the interview or self-completed questionnaire.

It must also be easily comprehensible, if subjects are going to be able to use it, and it should be encryptable, so that secrecy and confidentiality can be assured. A more important requirement is that the schema must be detailed enough and also open to modification; it must be complex enough to encompass all common (and more rare) sexual behaviour and situations, and allow the addition of new behaviour as it occurs.

The Inventory of Sexual Behaviour (ISB) was designed primarily for use in an interview situation, but was quickly adapted to use both as a self-administered instrument and as a Clinic-based checklist for use in taking a sexual history. Because it is based on the simple structure of the SBC (Sexual Behaviour Code) it is easily memorisable and can also be (and is) used to report sexual behaviour in covert contexts such as locations of casual sexual encounters.

Forms of the ISB vary from the simplest (designed to elicit information on anal intercourse only, which is 4 items long), through the basic heterosexual/homosexual behaviour (covering the four most common behaviours; 16 items long) and the fullest — so far! (which explores the 3 principal modalities for 6 main behaviours, looking at different genders of partners, condom use and differential behaviour between partners; this amounts to 720 items, but averages a completion time of only 18 minutes).

A question of terminology

The major question underlying ensuring the validity and viability of questions about sexual behaviour (which even in communities of sexual minorities are likely to be perceived as intrusive or at least delicate) is: 'How is sexual behaviour to be named?'. Such naming covers two different senses:

how is continuous sexual behaviour to be ‘chunked’ into recognisable and stable categories?

and

what are the categories to be called?

Fortunately for the anthropologist and survey researcher, the succession of (continuous) bodily movements that actually make up 'sexual behaviour' are, or can be, 'chunked' into identifiable and (well-nigh) universally recognizable sexual activity and given a common name. For instance, whilst the act of masturbation will usually have at least some unique components (for no-one does it exactly the same way, and no-one repeats the act identically), the manual stimulation of the penis by the hand is usually taken to be a necessary part of the definition. As in so many other aspects of sexual behaviour, even such a behaviourist as Kinsey insists that the definition must also have an intentional aspect (Kinsey, 1948:497–8) so that random, unintended, rubbing of the penis is excluded from the category of 'masturbation'. Yet recognition of the act of masturbation would be virtually universal, at least within a given culture, and although the exact temporal bounds (especially the exact point at which it may be said to begin) are rarely entirely agreed on, that the act occurred is universally acknowledged. Given that it is without doubt the most prevalent male sexual act, this is as it should be.

Despite its virtual universality and the early age at its first appearance, the nomenclature for masturbation is far from universal. Because it is often taboo among children, and frequently discouraged or prevented, it comes to be referred to by all sorts of euphemisms and code-names (often unique to the family concerned). There thus arises a hierarchy of terms of differing acceptability, from the medical terminology used by professionals (and often by subjects when talking to professionals) through a widely-used set of common vernacular terms to largely idiosyncratic ones. In SIGMA studies such terminology has been elicited before questioning detail of sexual behaviour. The purpose of this is not only to gather information on 'street' terminology, but
also to make the respondent more at ease in asking detailed information about what may be an embarrassing topic. In Question Schedule 1 of Project SIGMA, subjects were asked in the interview context to give their 'preferred name' for a range of common sexual terms. These were then substituted in questions using the terms. Although there was some differences in usage by areas (eg South Wales especially used terms like 'bonking' for (anal) intercourse), distributions were markedly similar.

**Preferred term for sexual terms**

- Preferred term for 'penis'
  - 53% XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX Cock
  - 18% XXXXXXXX Prick
  - 15% XXXXXXXX Penis
  - 8% XXXXX Dick
  - 5% XXXX Willy
  - 1% X Other

- Preferred term for 'semen'
  - 44% XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX Come/Cum
  - 38% XXXXXXXXXXXXXXXXXXXXXXXXXXXXX Spunk
  - 8% XXXX Sperm
  - 5% XX Semen
  - 6% XXXX Other

**Preferred terms for sexual behaviours**

- Preferred terms for 'masturbation'
  - 77% XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX Wank
  - 18% XXXXXXXX Masturbation
  - 4% XX Pull/Toss off
  - 1% X Other

- Preferred terms for 'fellatio' (oral sex)
  - 61% XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX Suck
  - 21% XXXXXXXXXXXXX Blow-job
  - 8% XXXX Fellatio
  - 8% XXXX Oral sex
  - 2% X Other

The full complexity of this approach to collecting information on sensitive sexual behaviour is described in (Coxon, 1988; 1992); suffice it to say here that it neatly combines the need to keep equivalence of meaning, ensuring compatibility and not unnecessarily embarrassing the respondent.

**Further strategies**

Finally, other strategies have been learned, taught by experience or tested out. They form an heterogeneous collection, so no attempt will be made to systematize them at this point.

*Co-operation with the Gay community* As a matter of principle the Project has worked directly with the gay community from the earliest planning stages of the Project, and most investigators and staff have been involved at some level in gay social, counselling or health organizations. The project has also reported back to the gay press, circulated a regular Newsletter, organised meetings in gay venues and talked at gay conferences. The Principal Investigators have all been up-front in avowing their sexual identity as gay men in all contexts, and have both interviewed and been interviewed as subjects themselves. Moreover, the Project has been able to give something back to the gay community by providing Hepatitis B vaccination free to subjects in the Project setting.

*Choice of interviewer* Respondents have from the outset been able to choose (or refuse) a given interviewer, either by name or in terms of their gender and/or sexual orientation. Although most research staff have been gay or bisexual, staff include heterosexual and lesbian women and heterosexual men. Although this allocation is of course far from random, there are no clear interviewer effects due to this factor. Interviewer recruitment, selection and training has been rigorous, and involves regular monitoring and de-selection where necessary.
One notable interviewer effect, due to gender, was that men interviewed by a woman were less likely to report vaginal intercourse than men interviewed by a man.

Debriefing procedures It became apparent that despite general willingness (and even enthusiasm) for a subject to share a good deal of intimate detail with the interviewer, there were certain areas of sensitivity, ‘shading’ and even downright lying. Impressionistically, these included minority sexual behaviour (eg fisting), anal-branchial insertion (‘fisting’), sexual practices (eg transvestism) and sexual pursuits (eg outdoor or toilet sex). Less expectedly, it also turned out that probably the most sensitive issue of all was sex for money (in either direction). This fact was discovered primarily in a de-briefing procedure where we concluded in lying by saying ‘Well, we know people sometimes lie through their teeth about some things in the interview . . . would you like to say where you did (lie)?’

Conclusions

In this sensitive area of research an especial debt is owed to the much-maligned influence of Kinsey and his team; there were situations when his question format: ‘When did you last . . . 1?’ (ie assuming the activity involved), was by far the best option. Even though co-operation has been maximised by methodological, strategic and other means outlined here, there is a strong relativity here: within a marginalised or socially invisible group the same issues arise as in the general population – there are still sensitive areas (but different ones) which require much the same techniques to overcome; it is simply that a conventional approach would be doubly insensitive. Estimates in these areas of sensitivity lead to prevalence and incidence figures which are systematically downwardly biased. Whilst Project estimates of (say) anal intercourse among gay men are very good ones which stand up to convergent validation, our estimates of male prostitution, casual sex, SM etc are biased downwards, and for this reason have mounted separate sub-projects to investigate them, using yet more innovative methods.

University of Essex

Received 14 April 1992
Accepted 24 May 1992

Notes

1 The experience of the Wellcome-funded UK study of general population sexual behaviour shows differences from the UK study in identifying those with homosexual experience.
2 Socio-sexual Investigations of Gay Men and AIDS. We are grateful for the (UK) Medical Research Council and the (UK) Department for Health for funding this study. The views expressed are however those of the Investigators and not necessarily those of these agencies.
3 Usually a blood sample, but also – alternatively or additionally – a saliva or (rarely) semen sample.
4 The exact area of the latter was the old county of Glamorgan, centred on Cardiff.
5 Permission had, of course, to be gained from area Ethical Committees, but in the South Wales sample there had been a long history of hostility between the gay community and the GUM Clinic, so it was imperative not to be seen to be associated with it.
6 In the case of London, this was directly under Dr McManus, but in Cardiff it was done under the authority of the University Occupational Health service. Since blood samples needed to be stored temporarily before dispatch for testing in London, a medical ‘hazard-room’ was designated and equipped. It is symptomatic of ‘AIDS hysteria’ that since the samples were potentially HIV antibody positive, this gave rise to fear among some academic – mostly social science – staff about the unlikely (indeed bizarre) conditions under which they might be at risk of infection, and in some cases led to formal complaints being laid against the Project to the University. Blood testing and training has been provided by the Virology Department at Dulwich Hospital, London, under the supervision of Dr Sheena Sutherland, whose help we gratefully acknowledge.
7 Several of the investigators and staff were (or had been) active in local gay community counselling organisations, and the health officers of the local Clinic (who themselves gave HIV test results) also helped in the training scheme. The arrangement was that all subjects who agreed to give a blood sample were pre-counsellled about the nature of the test and its results, whilst those who wished to know their results were given further counselling and allowed to change their mind about being told. They were also post-counselling by the Principal Investigator after the result had been given in a face-to-face session.
8 Blood tests results were given only to the Principal Investigator at the site concerned. He then called the subject in (whether the result was positive or negative) gave a resume of the meaning of each outcome, told him the result and counselled him about its implications. Blood test results were kept in a secret file in a separate micro-computer and could not be accessed by individual researchers. When it came to data analysis, a subfile was created which contained only the required variables, the subject identifier was stripped off, and the order of cases randomised. Blood test results were only given to the national (UK) Communicable Diseases Surveillance Centre in aggregate form.
9 There were a few female interviewers, and subjects were allowed to opt for either male or female interviewer but, as it happened, no female interviewers happened to be phlebotomists. Where they occur, gender terms are used descriptively and not generically in this paper.
This conflict has never been successfully resolved, and to the charge that it is likely to produce downwards estimates of risky behaviours, we can only reply that the multiple reports on sexual behaviour in the interview/s and sexual diaries, and some cross-referencing from sexual partners indicates that this does not often occur.

This read as follows:

**A CONFIDENTIALITY GUARANTEE**

Each volunteer taking part in the project is identified only by a code number when the data is analyzed. No individual is identifiable and published results will refer only to groups. Using physical precautions and 'encryption' we can guarantee that an individual's identity cannot be known.

In the event, we must regret making this a universal prescription on identification, since it meant that no record linkage could be made to sexual partners, and it was therefore impossible (normally) to estimate important epidemiological parameters (such as the number of partners of partners) and investigate contact networks. In the latest wave we have modified the undertaking, asking subjects to name partners and we then guarantee the confidentiality of that information.

Disposal for such infraction did occasionally take place, in conjunction with the University's disciplinary procedures.

Without identification it was not even possible to be sure that the interview scripts were genuine.

The question asks prospective applicants a variety of questions, which can be as vague (and presumptive) as 'Have you ever had a blood test?', or 'Have you ever been a patient at a Genito-Urinary Medicine Clinic?' as well as the more obvious 'Have you ever taken a test for HIV infection?'

The wording in quotation marks is that of the relevant legislation and confirms both its badly-briefed nature and striking resemblance to some US state's legislation.

This refers to Coxon 1988, which includes sexually explicit language. The link to Section 28 was that subscribers included educational institutions under local funding, whose trustees might claim that the article 'promoted' homosexual behaviour. The Editorial Board, strongly supported by the publisher, firmly resisted this threat and published the article in its entirety. Needless to say, no prosecution followed.

Sado-Masochist or 'Slave and Master' sexual activity.

This ruling was upheld on appeal to the House of Lords, though it is still subject to appeal to the European Court of Justice.

In Scotland, the restriction to only two men is not written into the law.

These included sexually explicit scenes of homosexual behaviour, used to check the reliability and concordance among interviewers of use of the Sexual Behaviour Code (Coxon 1988). After intervention by the Department of Health, the videos were released on an annual basis on signature of an official form entitled 'Release of Pornographic Material for Medical Research' [sic].

Some (unknown) person sent a xerox copy of the First Wave Question Schedule through the post to a man whose wife intercepted it, and referred it to the police. After negotiations with the Medical Research Council legal staff, no further action was taken when we were able to prove that we, the Project, had not sent it. Prima facie the Schedule, it appears, was 'pornographic' and hence was illegally being transmitted through the post.
Sex role separation in sexual diaries of homosexual men

Anthony P.M. Coxon, N. Huw Coxon with Peter Weatherburn, Andrew J. Hunt†, Ford Hickson, Peter M. Davies and Tom J. McManus*

Objective: To measure types of sex role prevalence in common and risk-related behaviours among gay men for modelling HIV transmission.

Design: Cohort study of 385 homosexually active men recording sexual diaries over 1-month periods.

Methods: Measures of incidence of behavioural sex roles for masturbation, fellatio, anal intercourse and anilingus by relationship type, derived from 1-month sexual diary data.

Results: Low behavioural role rigidity for masturbation and fellatio, but higher rigidity for anal intercourse and anilingus. Participants with no regular partner showed a relatively low frequency of anal intercourse, whereas those in closed relationships showed a high frequency.

Conclusion: Although anal intercourse shows a certain degree of behavioural role rigidity, this rigidity is not large enough to conclude that gay men exclusively engage in either an active or a passive role. Typical rates for exclusive active and passive roles for anal intercourse during the month the diaries were recorded were in the range of 12–15%; the dual role was significantly higher.

AIDS 1993, 7:877–882

Keywords: Homosexual behaviour, HIV transmission, modelling, behavioural role separation, sexual diaries.

Introduction

In the decade since the identification of HIV as the causative agent of AIDS much effort has been devoted to the prediction of transmission due to various environmental and behavioural factors [1]. Abramson [2] has noted that sexual parameters are critical to epidemiological models of the sexual transmission of HIV. The ultimate efficiency of such models depends upon quality and quantity of empirical data on sexual behaviour, predicting trends in the AIDS epidemic requires multidisciplinary co-operation [3].

The spread of HIV within the homosexual community is most closely associated with anal intercourse [4,5]. More specifically, passive anal intercourse is by far the highest risk factor in sexual transmission of HIV [6–9]. Conclusions derived from the disparity in risk between inserter (‘active’) and insertee (‘passive’) anal intercourse for epidemiological modelling were tentatively proposed inter alia by Trichopoloulou et al. [10], who suggested that exclusive behavioural role separation leads to a low incidence of HIV, citing the Middle East and southeastern Europe as examples. Basic models using this idea have been investigated by Wiley and Herschekorn’s [11], who confirmed Trichopoloulou’s conjecture. Wiley’s et al. modelling study also concluded that the ‘worst case’ epidemic occurs not with random mixing between groups, but when dual role individuals tend to mix among themselves.

The model of the structure of sexual action underlying the investigations of Project SIGMA (Socio-Sexual Investigations of Gay Men and AIDS) defines four possible modalities [12] of sexual behaviour: Self, Active, Passive and (simultaneously) Mutual. It is important to note that Active and Passive modalities are not always equivalent to who is the inserter and who is the insertee, but are in contrast directional (i.e., to who is the agent and who is the recipient of the sexual behaviour). The active partner is the inserter in anal intercourse, but the insertee for fellatio. This appar-

From Project SIGMA, Department of Sociology, University of Essex, Colchester and *Kings College Hospital, London, UK. † Who died from an AIDS-related illness during the period following submission.

Sponsorship: Supported by grants from the World Health Organization and the UK Department of Health.

Note: The views expressed are those of the authors and not necessarily those of the funding bodies.

Requests for reprints to: Prof. A.P.M. Coxon, Department of Sociology, The University of Essex, Wivenhoe Park, Wivenhoe, Colchester, Essex CO4 3SQ, UK.

Date of receipt: 4 November 1992; revised: 3 March 1993; accepted: 9 March 1993.
ent inconsistency is due to the fact that the modality is translated from the vernacular descriptions used by respondents (for example, 'I sucked my partner' and 'I fucked my partner' are both active behavioural roles, although the subject is inserter in the first and inserter in the second behaviour).

Four mutually exclusive behavioural role types can then be derived from these modalities: Both active and passive, only Active, only Passive and Neither (this partition of sex roles is referred to using the acronym BAPN). The 'Neither' role is crucial for accurately defining the incidence of a behaviour, since it makes the role set exhaustive. These behaviour role types are also fundamental in the sense that common, but more general, categories can be built up from them. Thus, 'inserter in anal intercourse' consists of two types: those who are 'Active' and those who are 'Both' in this terminology. More general categories can thus be matched by the BAPN behavioural role types.

The behavioural roles thus defined are best thought of as dispositional, but require a period of time in their definition, since at any one instant a respondent must assume either an Active, Passive or Neither role, but the individual's stability in a behavioural role is an issue of empirical fact. Thus, in a given period, interest centres on how consistent these modalities are; the Both role is expressed when an individual is both active and passive during this period.

The sexual diaries in this study were recorded over a period of 1 month, since this is the longest feasible period for collecting data by this method. Project SIGMA interview data indicate that estimates of the incidence of such behavioural role segregation for 1 month are highly correlated with estimates for a 6-month period, but the exclusive (Active, Passive) categories have a reduction in incidence during a full year. Such turnover is considered in the discussion of Table 3 below.

The incidence in each behavioural role for a population will therefore provide the necessary parameters for describing dominant HIV transmission routes, ignoring variations created by the use of prophylactics. In particular, we focus on the effects associated with relationship type and age, which Project SIGMA has found to be crucial independent variables [13].

Methods

Sexual diaries submitted to Project SIGMA were used as the data source for this study. The Project SIGMA diary study is of 385 non-clinic-based homosexually active men. These diarists are from various locations in England and Wales: primarily London and Cardiff, together with Newcastle, Teeside, Birmingham, Portsmouth, Leeds, Norwich, Liverpool and Bristol. Respondents from the diary study rarely participate in it for consecutive waves. The greatest overlap occurs between waves I and II, when 25 out of the 221 wave I individuals also recorded diaries in wave II.

The sexual diaries database relating to the 385 individual diarists comprises 5712 sexual sessions, consisting of a total of 12481 sexual acts. The median number of sessions per individual over the month is 15 (interquartile range, 11). The distribution of session length (number of acts in a session) is dominated by one-act sessions (over 70%), and follows an exponential decay function, with a mean of 1.90 and a variance of 3.06.

The validity of sexual diaries as a method for collecting sexual behaviour has been questioned [14], although James et al. [15] found a good fit between interviews and a self-administered questionnaire. However, Connath et al. [16] are unquestionably in favour of diary recording techniques. The SIGMA sexual diaries present an excellent methodology for recording detailed sexual information [17], especially about role expression, since respondents are explicit not only about what is done (the behaviour), but also to whom it was done (the modality, or role).

Project SIGMA has developed a model of sexual behaviour [12] that defines every sexual act as a composite of three elements: the modality ('who does what to whom'), the behaviour ('what is done') and the outcome ('the destination of ejaculate, if any'). This structure is used to encode the diaries into a database.

Each respondent completes a diary form that has two sections: the face-sheet with covering information, and the diary itself. The covering information includes demographic attributes, the SIGMA respondent type (derived from their relationship type and age; see below) and their HIV status (if tested). When completing the diary section, the respondent records details on where they were (for example, at home, in a sauna), when it was (the date and time) and who they were with (their sexual partners), and then describes, in sequence, the sexual acts that took place during the encounter. Since each sexual act is partner-specific it is possible to take the partner's characteristics, preferences (including power and behavioural role preferences) into account; the situational factors enable other issues of opportunity to be studied. Each diary is then transcribed into a database using the sexual code, ready for analysis by computer programs.

The respondent's age and relationship type are described using the SIGMA typology [13] (Table 1) and included with each individual's data in the database. To define the typology, the categories of sexual relationship type ('Closed' ('monogamous'), Open (at least one regular partner and others) or No regular partner) are crossed with the variable of Age (categorized as Under 21, 21 to 39, and Over 39 years). Relationship type is defined by the subject and may
be at odds with his partner's definition, particularly in the 'Closed' relationship. The youngest Age group consists of men whose sexual activity is illegal (in England and Wales), being under the age of homosexual consent of 21 years, and the oldest group consists of men who grew to sexual maturity before the 1967 Sexual Offences Act, when all homosexual activity (and hence, a fortiorti, male sexual relationships) was illegal. These nine SIGMA relationship types are then labelled using the Roman numerals I–IX.

Table 1. SIGMA typology.

<table>
<thead>
<tr>
<th>Relationship type</th>
<th>Age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Under 21</td>
</tr>
<tr>
<td>Closed (monogamous)</td>
<td>I</td>
</tr>
<tr>
<td>Open (one regular and other partners)</td>
<td>IV</td>
</tr>
<tr>
<td>No regular partner</td>
<td>VII</td>
</tr>
</tbody>
</table>

SIGMA, Socio-Sexual Investigations of Gay Men and AIDS.

We have written computer programs (N.H. Coxon et al., unpublished data) to perform several methods of data analysis on diaries. The program for BAPN analysis is chiefly used here. This program identifies each individual's data and assigns a role type for each behaviour during the specified period. The results are then aggregated into the nine SIGMA age x relationship types as percentage incidence figures.

Results

Table 2 shows the basic results from the BAPN program for the three most common behaviours (anal intercourse, fellatio and masturbation) and for anilingus (included because of its relative frequency and its implication in transmission of hepatitis A). The number of respondents in each type is shown under the type label. Each row of Table 2 represents a single behaviour-by-role combination, and each column a given type of diarist. The entries are the percentage of that type who engaged in the specified activity during the diary month. Each row was then analysed in a two-way table format (reflecting the structure of the 3 x 3 SIGMA age by relationship typology) using a median polish [18] (a resistant variant of the additive model). This method iteratively estimates main and interactive effects associated with Age and Relationship type categories on each sexual behaviour. The results of the analysis are not presented here, but the subsequent paragraphs rely upon them.

From the total (right-hand column) values in Table 2, it is clear that the two most common sexual behaviours (masturbation and fellatio) are little prone to behavioural role rigidity, in the sense that the 'Both' category dominates the others and the exclusively Active and exclusively Passive roles are small. In contrast, anal intercourse shows a considerably higher degree of role rigidity, and anilingus a small (but significant) degree of rigidity.

On analysis of the four rows of the Anal intercourse subtable (Table 2), a consistent pattern of effects emerged. In general, Relationship type had stronger effects than age on the incidence figures of each variant of anal intercourse. In particular, having No regular partner lowered the incidence of anal intercourse for all practising (Both, Active or Passive) roles, and being in a Closed relationship consistently raised it. Age effects were concentrated mainly in the youngest group (under 21 years), considerably decreasing the incidence of anal intercourse for those exclusively Active and increasing the incidence of those exclusively Passive. A few strong interaction effects augment this picture — the incidence of Active anal intercourse for those under 21 years and with no regular partner was much higher than would be expected given the main effects, and the incidence for those in a Closed relationship and under 21 years was much lower than expected. In brief, the Regular/No regular partner divide and the Under 21/Over 21 years boundaries are the most significant with respect to sexual role differentiation according to this analysis.

In contrast to anal intercourse, there is no real common consensus amongst researchers on the level of risk associated with fellatio. Analysis of the sexual diary data of which sexual behaviours immediately precede anal intercourse in a sexual session has shown that fellatio is by far the most common antecedent [19],
complicating measurement of its independent level of risk. However, cases of HIV transmission via the oral-genital route have been reported [20], making fellatio a possibly significant component of a transmission model, especially considering the high infectivity for fellatio during primary infection (523 times higher than at any other stage of infection) reported by Kopperman et al. [21]. Analysis of this section of Table 2 shows that respondents expressing the Both role in fellatio occurred mainly among those in an open relationship, and were fewest among those with no regular partner. The Active role was associated most strongly with those over 39 years old, and least in Closed relationships. The Passive role was highest in Closed relationships, and lowest for young men (under 21 years). As mentioned above, 44% did not engage in fellatio during their diary. The Neither role was strongest among those under 21 years old.

Masturbation is a low-risk behaviour, but is useful as a 'yardstick' because it is common. Analysis of this section of Table 2 indicates that Age and Relationship effects were very small for masturbation, especially among those who practised the Both role (67%). It is likely that, given a few more months, almost every respondent who engages in masturbation will do so in both roles.

The HIV transmission risk associated with analingus is not well documented, but it is still an important variable in the transmission of hepatitis A and associated sexually transmitted diseases. Analysis of this section of Table 2 shows that the Active role had the highest incidence for those in an Open relationship, and lowest for those in a Closed relationship. The highest incidence for Passive anilingus was associated with men between 21 and 39 years old. Those who engaged in both roles did so the most in a Closed relationship (a similar pattern to anal intercourse).

How stable are these role modalities? To answer this question, turnover data are necessary, and unfortunately the longitudinal properties of sexual diaries data in Project SIGMA are poor: relatively few men keep a diary on a regular basis, and since only 25 individuals wrote diaries in both waves of these data, it is hazardous to rely on their turnover rates. However, these 25 diarists are a subset of the 296 SIGMA respondents for whom we have reported (interview) sexual behaviour data on an annual basis. Therefore the turnover rates between BAPN role types based upon the interview data of the Project for this year are shown in Table 3.

Table 3 shows the turnover rates (in number of individuals) for each of the role variants of anal intercourse. The intersection of a row with a column shows the number of individuals who were in the row's role in wave III who changed to the column's role in wave IV. The diagonal entries show the number of individuals who remained in the given behavioural role in both waves. The most striking feature is the high stability of the Both and Neither roles: approximately two-thirds of those who abstained from anal intercourse (Neither) and of those who practised it both actively and passively (Both) continued to do so a year later. Thus, the Both category is the largest and the most stable role behaviour among those engaging in anal intercourse. Those who are exclusively Active or exclusively Passive not only form a much smaller proportion of gay men, but were almost twice as likely to change to another role. A number of international studies reported by Ross [22] have shown that preferences for sexual acts in their various modalities also follow a similar pattern, with the Both variant of anal intercourse being considerably more common than the single-mode Active and Passive variants.

Finally, we consider the individual-based average frequencies of the behaviours during the diary/month (derived by counting their occurrence in the diary records), since such measures are also crucial parameters for modelling sexual behaviour.

The four most frequent sexual activities shown in Table 4 (masturbation, fellatio, anal intercourse and anilingus) account for approximately 80% of total non-solo sexual activity (i.e., with a partner present), and together with their reciprocated modes (mutual masturbation and mutual fellatio) represent over 90% of the total number of sexual acts. The mean frequency rates per individual for the four behaviours are also shown in Table 4. The mean number of sexual acts per individual diary/month is 19.3 for these behaviours.

On average, anal intercourse was engaged in three times a month. Since over 50% of diarists did not engage in anal intercourse (this is a typical value for the Neither role; see Table 4), those who did engage in anal intercourse did so approximately six times a month. The mean values for other behaviours can similarly be adjusted by reference to the appropriate 'Neither' category in Table 3.

### Discussion

Men who engaged in anal intercourse did so a mean of six times a month, out of a total of approximately 20
Table 4. Mean incidence of behaviours per diary.

<table>
<thead>
<tr>
<th></th>
<th>Mean incidence per individual (diary)</th>
<th>Total for behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anal intercourse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>1.6</td>
<td>2.9</td>
</tr>
<tr>
<td>Passive</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>Fellatio</td>
<td>2.9</td>
<td>5.2</td>
</tr>
<tr>
<td>Active</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>Passive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masturbation</td>
<td>5.2</td>
<td>9.5</td>
</tr>
<tr>
<td>Active</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>Passive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anilingus</td>
<td>1.0</td>
<td>1.7</td>
</tr>
<tr>
<td>Active</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>Passive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>19.3</td>
</tr>
</tbody>
</table>

sexual acts a month. The monthly incidence of Both, Active, Passive and Neither roles for anal intercourse was 18, 14, 12 and 56%, respectively. When broken down by Relationship type and Age group, the highest levels of the Both role were among those in a relationship (whether Closed or Open) and by far the lowest incidence of anal intercourse was among men with No regular partner. Similarly, the Both role was particularly common in the 21–39 years age group.

The lowest incidence of anal intercourse was associated with men with No regular partner (who had the most casual encounters in the diary data). Approximately 72% of them did not engage in anal intercourse in that month. The highest incidence occurred in Closed relationships. This suggests that frequency of anal intercourse is inversely related to number of sexual partners, which further supports the need to use the variable of penetrative sexual partners [5] in epidemiological models (since the number of sexual partners is not indicative of the number of risk contacts).

These results therefore strongly contradict the idea that homosexual men tend to be exclusively active or passive during anal intercourse in England and Wales; rather, if within a given period they engage in anal intercourse they are overwhelmingly likely to do so in both modalities. Indeed, the figures we have cited may well be an underestimate [10], especially since the number of respondents practising the Both role is less than figures observed by Van Zessen and Van Griensven [23], who reported an incidence of 61% for the Both role in their Amsterdam cohort (although this is over a period of 6 months using retrospective interviewing techniques). However, our diary-based BAPN estimates for anal intercourse in England and Wales have been confirmed independently by retrospective interviewing results from the London Project SIGMA site, who report interview-based Both, Active, Passive and Neither rates of 17, 14, 12 and 57%, respectively, over 1 month (P. Weatherburn, personal communication). These figures agree with ours within a range of 1%.

Wiley and Herschekorn’s [11] epidemiological analysis of homosexual behavioural role separation in anal intercourse used the simplifying assumption that role characteristics are constant over time (i.e., that individuals maintain their sex role type consistently). Clearly this is not likely. Our turnover table (Table 3) suggests that while 65% of men who engage in Both modalities are likely to maintain their role over a year, only a third of those exclusively Active (39%) or exclusively Passive (34%) are likely to do so. In their study of behavioural role change rates in a large longitudinal cohort (n = 723) in Amsterdam Van Zessen and Van Griensven [23] reported that roles during anal intercourse are significantly static over time: the percentage of individuals who maintained their sex role type on follow-up after 6 months was 62, 42, 29 and 78% for Both, Active, Passive and Neither, respectively — which corresponds within ±5% to the SIGMA interview-based annual rate in Table 3. They also reported a similarly low exchange rate (‘migration’) between the Active and Passive roles for anal intercourse.

The characteristics associated with risk behaviours presented here have important implications for modelling of HIV transmission in the homosexual population.

First, if gay men engage in anal intercourse they are far more likely to practise both Active and Passive variants within a given period than to practise one role exclusively.

Second, as Wiley et al. [11] conclude, the worst case epidemic occurs when individuals expressing the Both role tend to mix among themselves. This diary study therefore clearly pinpoints gay men in a Closed relationship and men in the 21–39-year age group as potentially exhibiting the highest risk of HIV transmission, due to the dual-role mixing level associated particularly with them. This is a worrying conclusion given that Regular partnership durations are often shorter than the period of primary infection (i.e., before an effective antibody response is developed). Moreover, since few gay men in England and Wales are normally tested more than once a year, this duration is also smaller than the latency between tests.

Third, although young gay men (under 21 years) are generally slightly more likely to engage in anal intercourse (especially passive anal intercourse), they are much less likely to do so in the dual role.

References

SHORT COMMUNICATION

No connection between alcohol use and unsafe sex among gay and bisexual men

Peter Weatherburn, Peter M. Davies*, Ford C.L. Hickson, Andrew J. Hunt, Thomas J. McManus† and Anthony P.M. Coxon*

Objective: To investigate the relationship between alcohol use and unsafe sexual behaviour.

Methods: The paper discusses data collected from 461 gay and bisexual men interviewed in England and Wales by Project SIGMA during 1991-1992. These data were collected during face-to-face interviews using retrospective weekly diary techniques and include details of all sexual sessions and alcohol use. The 819 reported sexual sessions with other men are divided into those that involved alcohol use (30.6%) and those that did not.

Results: Differences in the incidence of HIV risk behaviours between sexual sessions that involved alcohol use and those that did not are small, and none are statistically significant. Furthermore, for those men who engaged in sexual behaviour whilst under the influence of alcohol, the quantity of alcohol consumed had no effect on sexual behaviour.

Conclusions: Among gay and bisexual men, sex under the influence of alcohol is no more likely to be unsafe than sex among men who have not consumed alcohol.

AIDS 1993. 7:115-119

Keywords: Gay men, sexual behaviour, alcohol use, unsafe sex, anal intercourse, condom use.

Introduction

Since 1986, a vast number of papers examining the relationship between alcohol use and unsafe sexual behaviour have been published (see [1,2] for reviews). As a result, many health educators and researchers appear to assume that it has been proven that alcohol consumption leads to an increased probability of engagement in unsafe sexual behaviour.

However, this literature contains a number of major limitations. The first arises as a result of the widespread assumption that there is, or could be, only one straightforward relationship between alcohol use and sexual behaviour, irrespective of factors such as culture, gender or sexual orientation. Thus, despite evidence that the relationship varies according to sex and sexual orientation [3,4], results from samples of gay and straight men and women are frequently considered together.

Second, the literature is littered with methodological limitations. The most important is the conflation of alcohol and other drugs (including such diverse substances as marijuana, cocaine, and 'poppers') in one global measure [3,5-8] assuming that all drugs are used in similar circumstances and to obtain similar effects. Another problem arises with the comparison of past sexual behaviour with gross measures of previous alcohol consumption (typically drinking sessions per time period) [5,8-22]. Since both variables represent gross measures of past behaviour any association tells us little about the actual relationship, since not only are there likely to be many intervening factors.
but the alcohol use need never have coincided with the unsafe behaviour [23].

However, even if we choose to overlook these methodological shortcomings, examination of the literature pertaining to gay and bisexual men shows that evidence for any relationship is at best tentative. From our extensive review, only eight papers [3,5–7,9–11,24] document a significant relationship. 15 [1,8,12–19,25–29] fail to do so and a further four are equivocal [4,20–22]. Furthermore, all those that use critical incident [17,25–27] or diary techniques [29] to test the relationship between alcohol use and unsafe sex fail to find any association.

Methods

Cross-sectional data presented refer to the 461 respondents interviewed by Project SIGMA (Sociosexual Investigations of Gay Men and AIDS) in the fourth wave (1991–1992) of a 6-year, five-wave cohort study of homosexually active men. The original cohort was recruited in 1987/1988 by a variety of means, including a postal questionnaire in the gay press; recruitment in gay pubs, clubs and social and political organizations; and contacts of the above. No respondents were recruited from genito-urinary medicine clinics. Respondents live in and around 10 main sites across England and Wales. An extensive attrition analysis [30] reveals that the only factor that differentiates responders and non-responders is age: younger men are less likely to respond. However, wave 1 and wave 4 samples are almost identical demographically due to careful recruitment of new respondents in wave 4.

During the interview retrospective weekly diary data were collected from all respondents. Data on the incidence of all sexual behaviour and alcohol use in the preceding week were collected in a diary format noting, for sexual sessions, the day, time, partner, place and exact sequence of sexual acts including use of condoms and lubricants, and, for all alcohol use, the day, time, units, and perceived effect of alcohol intake. Sexual diary entries were returned by the interviewers using shorthand codes designed and widely used by Project SIGMA for prospective monthly sexual diaries [31].

The main units of analysis were sessions (drinking and sexual) and not individuals, although we still rely upon the reporting of one individual in any sexual dyad. The main advantages of this method are that, although we can still count the incidence of activities, sexual behaviours are contextualized and patterns in the interaction of activities can be observed.

Sexual sessions were divided into those that involved alcohol before and/or during, and those that did not. To qualify as a alcohol and sex session, the hours elapsed from the end of alcohol intake until the beginning of the sexual contact had to be less than the total number of units consumed. Based on the assumption that alcohol is 'cleared' from the body at an average rate of 1 unit per hour, this convention was designed to minimize subjective judgements. Surprisingly, there were very few instances where it was necessary to apply this convention. The vast majority of sessions included in the alcohol and sex category involved sexual activity initiated less than an hour after alcohol consumption had ceased.

Given that the sexual behaviour of gay and bisexual men is contingent upon the type of partner (casual or regular) involved [32], we also distinguish between sessions by partner type. We define 'regular partner' as a partner with whom you have had sex more than once, where the second and subsequent meetings were not accidental, and with who you intended to have sex in the near future. The term 'casual partner' is used to describe partners who fall outside the definition of regular sexual partner.

Results

In the week preceding interview the 461 respondents reported 1624 alcohol use sessions; 1208 instances of self-masturbation; and 841 dyadic sex sessions. Of these 841, 33.5% (n = 282) were with casual male partners, 65.9% (n = 537) with regular male partners, and 2.6% (n = 22) with a regular female partner. In the following we refer only to sexual sessions with other men (n = 819).

Anal intercourse and condom use

Overall, 24.8% (n = 203) of sexual sessions involved anal intercourse. That is, 10.7% (n = 88) included only receptive, 12.6% (n = 103) only insertive, and 1.7% (n = 14) both receptive and insertive. Since engagement in both modalities of anal intercourse in the same sexual session is relatively rare, such incidents are included in the figures for both these activities.

Table 1. Percentage of sessions involving anal intercourse and percentage of those in which condoms were used.

<table>
<thead>
<tr>
<th>Sexual activity (%)</th>
<th>Overall (n = 819)</th>
<th>Regular partners (n = 537)</th>
<th>Casual partners (n = 282)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Receptive</td>
<td>Insertive</td>
<td>Receptive</td>
</tr>
<tr>
<td>Anal intercourse</td>
<td>12.5</td>
<td>14.3</td>
<td>15.3</td>
</tr>
<tr>
<td>Condom use for anal intercourse</td>
<td>46.1</td>
<td>51.3</td>
<td>39.0</td>
</tr>
</tbody>
</table>

As Table 1 shows, sex sessions with regular partners were more likely than those with casual partners to involve both receptive ($\chi^2 = 11.3$; d.f. = 1).
Alcohol use and unsafe sex  Weatherburn et al. 117

In summary, the differences between those sex sessions that took place under the influence of alcohol and those that did not were small, and none were statistically significant.

**Quantity of alcohol consumed**

Table 4 shows the average quantity of alcohol consumed, per drinking session, across the range of markers of unsafe sex we have used above.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Median [IQR]</th>
<th>F ratio (d.f.)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All drinking sessions</td>
<td>4.5 [2.1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Yes: 5.5 [6.3]</td>
<td>32.001(1,1610)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>No: 4.4 [8.1]</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Partner</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Casual</td>
<td>6.7 [3.1]</td>
<td>5.151(1,1264)</td>
<td>&lt;0.03</td>
</tr>
<tr>
<td>Regular</td>
<td>4.5 [9.4]</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Anal intercourse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>6.6 [6.2]</td>
<td>0.96(1,2441)</td>
<td>0.34</td>
</tr>
<tr>
<td>No</td>
<td>5.6 [6.2]</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Condom use for anal intercourse</td>
<td>6.6 [6.2]</td>
<td>0.005(1,56)</td>
<td>0.95</td>
</tr>
</tbody>
</table>

IQR, interquartile range.

Having removed far outliers (data that lie three interquartile ranges above the upper quartile) [33], we found that men who drank and had sex consumed significantly more units of alcohol than men who just drank, and men who had sex with casual partners consumed significantly more than men who had sex with regular partners. However, men who engaged in anal intercourse consumed no more alcohol than those who did not, and those who used condoms for anal intercourse consumed no less than men who did not.

**Discussion**

The data we have described show no statistically distinguishable differences in HIV risk behaviour between sessions that take place under the influence of alcohol and those that do not. Gay and bisexual men in this study were no more likely to engage in any of a range of risk behaviours after consuming alcohol than otherwise, and for those who did have sex whilst under the influence of alcohol the quantity of alcohol consumed had no effect on risk behaviour. These results add to the body of evidence that fails to confirm the widespread belief that there is a direct relationship between alcohol use and engagement in unsafe sex.

In the first and, in retrospect, the most influential contribution to this debate, Stall et al. [24] found a relationship between drug and alcohol use and unsafe
sex. Whilst alcohol had the weakest effect on sexual behaviour of all the drugs reported, their discussion suggested it might be of central importance in understanding risk behaviours. Hence, they postulated a number of hypotheses that might account for this relationship, including 'aphrodisiac'; 'personality'; 'social context'; 'multifactorial'; and 'disinhibition'.

For any of these substantive hypotheses to be true, it must logically be the case that a relationship should pertain at the level of the individual session. An association at the gross level without an association at the basic level is strong presumptive evidence for a spurious effect in the former case. Our data add to a growing list of reports that have failed to find any relationship at this primary level. While, technically, our findings do not disprove the contention that alcohol consumption leads to unsafe sex, they provide strong evidence that the relationship found in early papers might be spurious. As we have argued elsewhere [23], this is most probably due to methodological inadequacies, specifically the use of gross measures of alcohol consumption and sexual behaviour.

Whilst earlier contributors to these debates should not be held responsible for the ways in which educators and policy makers have interpreted their works, early papers certainly contributed towards alcohol being seen as of central importance to HIV prevention campaigns.

Thus, health education campaigns have emphasized the disinhibiting effect of alcohol by stressing that 'when we get high, we are likely to be tempted into sexual activities which are riskier for AIDS' [34]. These campaigns not only make unjustifiable assumptions about the effects of alcohol on self-control, but also include comments on other drugs, invariably confusing their effects.

While our qualitative work [23] suggests that alcohol is an important factor in the sexual lifestyle of more than a fifth of gay and bisexual men, it is usually used to facilitate the social and sexual lifestyles these men have chosen to pursue. Most generally, alcohol is often used in a premeditated way, to enhance sexual desire or performance and/or overcome sexual and social inhibitions. Finally, some gay men stress that they use it to enable them to engage in sexual negotiations that they might otherwise be too shy to attempt. Crucially, these include the negotiation of safer sex. For these men, admittedly a minority, the information that alcohol use may impede their ability to have safer sex is clearly untenable.

We suggest that forceful messages suggesting that alcohol inhibits the ability to have safer sex may have little relevance to the majority of gay men, and that campaigns that posit a straightforward association between alcohol use and unsafe sex may be providing individuals with an easy self-justification for engaging in such activities. At worst, such campaigns may facilitate engagement in unsafe sex by allowing individuals to abrogate responsibility for behaviours that they or their peers find regrettable. Furthermore, by providing a ready-made agenda for intervention they absolve health educators from addressing the complex and intractable problems of negotiation and interaction that are central to any understanding of unsafe sex.

Clearly, we feel it is necessary to re-open debate on the topic of alcohol use and unsafe sexual behaviour. Future research could do much to clarify whether any relationship exists and to explain its implications if it does. However, further research should concentrate on methodologies that can examine the effects of alcohol (and other individual drugs) within the specific context of sexual encounters. In order to do so, methods must not assume that any psychoactive drugs have identical effects, or that any population group uses them for the same reasons, or in the same circumstances.

References


29. Leigh BC. Alcohol consumption and sexual activity as reported with a prospective diary technique. J Abnorm Psychol (in press).


Chapter 8

Diaries and Sexual Behaviour; The Use of Sexual Diaries as Method and Substance in Researching Gay Men’s Response to HIV/AIDS

Anthony P.M. Coxon

The Human Immunodeficiency Virus (HIV) is transmitted primarily in sexual activity; hence detailed information is needed about the sexual context in which that transmission occurs if accurate epidemiological estimates are to be made and the spread of the pandemic to be understood and contained. How such information may be obtained, and how its reliability and validity may be assessed form the basis of the research reported in this chapter. Very little can be taken for granted in such research and the methodological problems are considerable.

From the outset we know that sexual behaviour such as vaginal and anal intercourse are particularly implicated in transmission, but the evidence is much less clear for oral sex and other sexual activities and it is therefore necessary to investigate the whole range of activities which could possibly lead to transmission.

The individual is at the centre of that investigation and is normally the source of such data, so it makes sense to collect data from his or her (ego’s) perspective. But since any implicated sexual activity involves at least one other person, the real unit of investigation is the dyad (couple), and the behaviour and characteristics of the other person (alter) are also crucial.

Information of a particularly intrusive kind is needed for analysing HIV transmission. Since the mechanism involves the ‘exchange of bodily fluids’ — especially semen, blood and saliva — they too must be tracked and reported and details of whether and how ejaculation of semen occurs is an integral part of the analysis. Our research focus does not, however, cover all types of sexual behaviour, it is restricted to the sexual and social lifestyles of men who have sex with men¹ (and possibly with women).

If we were interested primarily in the actual mechanisms of sexual transmission of HIV, then direct observation might well be the appropriate method for obtaining information, as in Masters and Johnson’s (1966) work. But such a method is likely to lead to highly biased estimates, since only a highly
atypical sub-population is likely to consent, and the presence of an observer would itself be highly reactive. Observation as a method is therefore far from unobtrusive and would involve major problems of consent and organization (and cost). It would also be illegal, at least in England and Wales (but not in Scotland).2

In order to obtain relevant information, we have to rely on subjects’ own reports or accounts of sexual activity have to be obtained, and the interview setting provides the most usual context of data-collection. In Project SIGMA the yearly Core Question Schedule includes as a central element the Inventory of Sexual Behaviour (ISB) (Coxon et al., 1992b), asking respondents a systematic set of questions about whether (and if so, how often) they had engaged in these detailed activities (for prevalence) and within a given period of time (for incidence). But how accurate are such subjects’ estimates likely to be? From the outset of our enquiries the data gave good grounds for scepticism: the numbers given in answers were often suspiciously vague, rounded or approximate (Coxon, 1988b), suggesting problems of accurate recall. Moreover, when questions (identical or implied) were repeated later in the interview, the number given was rarely the same, suggesting problems of reliability. When cross-checks were made with estimates given by their partners the numbers were (to varying degrees) often at variance with the respondent’s, suggesting problems of validity.

Much information obtained about sexual activity is also atomistic and out of context — we learn whether or how often something was done, but rarely the context in which it took place, the sequence in which it occurred, or the person with whom it occurred. Such factors make a big difference to the meaning of sexual behaviour, but they are also important, we believe, in attempting to understand sexual risk. That sexually risky behaviour takes place is important, but if people are to be encouraged to lessen or avoid risk then we need to know the significance of such behaviour to the person, and we also need to identify its context in order to find out whether risk-taking varies systematically by situation, rather than simply by individual. Again, the number of sexual partners a person has is an important variable epidemiologically, but it is even more important to know whether they are one-off or regular partners, whether sex with such partners involves penetrative or unprotected (risks) sex, whether alcohol or drugs such as nitrates have been used (possibly as disinhibitors) … and so on. In the normal way such questions are asked separately, and even if recall is excellent we can know nothing about how they co-occur with sexual behaviour and combine in a particular sexual situation to increase or decrease risk. Finally, the order in which sexual activity occurs (and the position in a sequence in which an act occurs) can have quite different effects. An example is the differing risks of hepatitis infection when oral sex follows anal intercourse as opposed to preceding it. Similarly, the probabilities of transmission are very different according to whether a person is anally receptive or insertive, and we knew little indeed about the prevalence and possible mixture of sexual role playing in male-with-male sex (Coxon and Coxon, 1993a).

The interaction of these issues is highly complex, and answers to them cannot be obtained by simple questionnaire methods. But they are pressing issues, whose answers could have radically different consequences for understanding and predicting the spread of the pandemic and for health education and interventions. It is not just that recalling complex behaviour is more difficult than recalling simple behaviour but that many people are unaware of how these factors combine in their own case, or simply cannot give a verbal account of it. A rather different method is therefore called for which can provide information in a manageable and a systematic way: we developed the diary method for this purpose.

Diary-keeping is a very natural way to elicit data of the sort we require. Almost everyone has kept a diary at some point and it is often the chosen way to confide and record one’s thoughts and actions. It is, as Plummer (1983) rightly dubs it, a ‘document of life’. Used as a social science method, it can be a valuable non-reactive method (when previously written, uncommissioned diaries are used as a resource), or it can be a specially elicited record, typically focused on one domain such as purchases or alcohol consumption or, as here, on sexual activity. Like its natural variant, it suffers similar problems of motivation; many start diaries but fewer finish them.

How does the diary method compare with other methods? As in Content Analysis, diaries are usually written in natural language format, and are subject to the same forms of analysis including syntactic, semantic and thematic analysis. But in diaries, the focus and domain of interest is usually narrower than in most naturally-occurring prose. Like the similarly methods of Life-Course or Event-history, the diary method is time-structured and sequential, but it is usually more detailed and discursive in content and has a much smaller time-span than a history. There are even some similarities with the questionnaire. The questionnaire can differ in how structured it is and whether it is self-completed or administered in an interview situation, and so can the diary method. The main difference is that diary data are not elicited in a pre-ordained, conditional branching sequence of questions, as in the case of the questionnaire.

The Sexual Diary

The diary method has been developed within Project SIGMA (Coxon, 1988a; Davies and Coxon, 1990; Coxon et al., 1992b, Coxon and Coxon, 1993b) as a parallel to the Project’s more conventional methods, and has now become the preferred (indeed, the unique) method for obtaining certain sorts of information about the detail of gay men’s sexual activity. It is important to enter a series of provisos at this point:

(i) We are restricting attention here to sexual behaviour, other methods are used to establish the meaning and context of that behaviour and other
forms of data collection are useful in relation to other research questions.

(ii) The need to concentrate on sexual behaviour capable of leading to HIV transmission means that considerable (perhaps undue) attention is paid to ejaculation and its sequela. This aspect can, of course, be ignored and omitted.

(iii) Although the sexual diary method has been developed in the context of studying homosexual behaviour there is nothing to restrict it to this orientation. Indeed, sexual activity between homosexual men and their female partners forms a natural part of this study.

Project SIGMA has developed a theory (or schema) for the representation and analysis of sexual activity (Coxon et al., 1992b), which forms the basis for obtaining systematic information about the prevalence and incidence of sexual behaviour in both the Interview/Questionnaire context and that of the sexual diary (see Davies, Chapter 4, this volume). This makes aggregation and comparability of data from these different data sources a straightforward matter, and it also gives SIGMA’s use of the diary method a very distinctive flavour. The theory itself arose both as a way of systematizing and inter-relating the components of sexual behaviour relevant to HIV transmission and as an attempt to connect the structure of sexual behaviour to Talcott Parsons’ account of the Unit Act in the structure of social action, and to communication processes. At an early stage we had realized that the structure of sexual behaviour has a striking resemblance to linguistic structure, and that to interpret it in this way gives added insight to the analysis and meaning of sexual behaviour. In this interpretation, the self-contained unit of communication analogous to the sentence is the sexual session; the constituent words correspond to the sexual acts, and the inflections of the word could encode the activity, the modality and the outcome. The information transmitted in a sexual session is basically a predication of the form:

WHO does WHAT, TO WHOM and with WHAT OUTCOME

which encodes the agent (who), the sexual behaviour (what), the other sexual recipient (to whom) and whether and how ejaculation occurred (what outcome).

This same structure can then be used to define question-forms for questionnaires (such as the ISB) and for diary instructions and makes it possible to compare data having many different formats.

Advantages and Disadvantages

Before proceeding with a specification of how the method of diaries is applied to sexual behaviour, it is worth pausing to summarize the advantages and disadvantages of the sexual diary method. The sexual diary method is a more ‘natural’ method than most, both in the sense that it exists as a common social practice and that it is written in natural language. The diary makes it possible to obtain information in far greater detail than other methods, since it is designed to minimize recall and memory errors and cognitive strain. It is especially adapted to gathering reliable information on the time-sequence of events, so that change is easily charted. The information can be obtained in a contextually-specific manner, without relying on recall; thus variation due to such things as particular partners or particular settings can be directly studied. Quantitative information is derived directly from the data, without recourse to the errorful estimating procedures used by survey questioning or respondent recall. The sexual diary can be augmented to obtain other concurrent information such as alcohol and drug use in sex (see Weatherburn et al., 1993), and the data obtained are, on present evidence, more reliable than those obtained from retrospective recall in surveys (see Janson, 1990). These advantages are impressive, but need to be balanced against the undoubted disadvantages, some of which can be ameliorated.

The main disadvantages of the sexual diary method have to do with bias in recruitment of respondents rather than with the method of data collection per se. But there is undoubted selection bias with respect to those who do and those who do not agree to be diary respondents or return information. In the case of hidden populations like gay men (SIGMA, 1990; Coxon and Joyce, 1993) selection bias in the recruitment of those prepared to keep a diary exists in addition to that in the initial sampling procedure of Project sample members. The sources of bias are very similar to those in other studies relying on volunteer subjects (Rosenthal and Rosnow, 1975:225) who tend to be educated, of higher social class, intelligent, approval-motivated and sociable; (the last characteristic takes the form of being more likely to be 'out' as gay men). Those volunteering tend to be more sexually active (in the sense of having more sexual sessions and more partners) than those who do not volunteer. For longitudinal studies there is an undoubted ‘step-wise attrition’ — it is far easier to persuade men to keep a diary for consecutive months than regularly on a yearly basis. The type of data generated cannot readily by analysed by conventional packages, and rely on an intermediate stage of string-manipulation software (see section entitled ‘Sexual diary data’ below).

Implementation of the sexual diary

Keeping a sexual diary is not a novelty to many gay men (Joe Orton is a notorious instance; Lahr, 1986); some have been keeping one intermittently all their life. Sometimes this is instrumental (in case of infection with a sexually transmitted disease, so that partners can be informed), but more often it is intrinsically interesting, especially to those with a full and/or complex sex life, or
who wish to note their adventures for later enjoyment or as an aide-memoire to masturbation.

At each wave of the Project, the SIGMA respondent is taken through the last week of his sexual activity by the interviewer according to a specified format. Originally this was done to ensure that the respondent understood the instructions for keeping the diary. The interviewer often assisted in its recall and wrote the actual transcript. Subsequently, this procedure also provided useful information on autobiographical memory: how far back could he recall the detail of sexual activity? At the end of the interview, a month-long diary and set of instructions (see SIGMA, 1993) were given to the respondent for completion and return (an example of a completed week of such a diary is contained in Appendix One).

Respondents are told that accounts of sexual activity should be written down as soon as possible, and if possible on the same day they occurred; only to complete a diary if they are prepared to be completely honest, and not to invent or ‘shade’ what they do; that the basic unit of their account should be the ‘session’ (‘one or more sexual acts by yourself or with the same person/s at one time’). The format of each session, derived from the ‘structure of sexual action’ (see Coxon et al., 1992b) is then explained. The components are:

- **Time, place and antecedents:** Day, hour; location (for example, in whose accommodation or external sites, such as parks, toilets, the activity took place), together with antecedents such as the use of alcohol, drugs and nitrates;
- **the participants:** (if any); description of the sexual partner/s involved in this session;
- **the sexual activity:** for each constituent sexual act in a session — the behaviour; the modality (who did it to whom) and the outcome (whether and how ejaculation occurred);
- **accompaniments:** especially use of condoms, lubricants, ‘toys’, etc.

The relevance of the participants and sexual activity components has already been explained above, but the inclusion of the others needs a brief explanation. **Time** is necessary for sequencing sessions during a day (what one does sexually in the morning often differs dramatically from what happens in the evening). **Place** or location allows us to separate out home-based from casual or out-of-doors activity. **Antecedents** are part of the scene-setting which makes it possible to inspect the effect of precursor activities (poppers or alcohol before, as opposed to within sexual activity); together they define the situation and hence reduce problems of indexicality. **Accompaniments**, by contrast are part of the scene and sexual activity which do not usually have a direct effect of the probability of transmission.

---

**The Sexual Act**

The sexual act forms the basic building block — the word of the session’s sentence — and the structure here is crucial for later analysis. In SIGMA’s schema, all description is ego-centric, i.e., viewed from the respondent’s (ego’s) position in the proceedings, and the question of which sexual actor is doing a given act to the other is dealt with by the relational modality of the act. Thus, Active always means that ego does x to alter, and Passive always means that alter does behaviour x to ego. A simplified form of the specification of the sexual act is as follows:

\[
\langle ACT\rangle :: = \langle MODE\rangle \langle BEHAVIOUR\rangle \langle EGO'S ORGASM\rangle \langle ALTER'S ORGASM\rangle \langle MODIFIER\rangle
\]

\[
\langle MODE\rangle :: = [S|A|P|M]
\]

\[
\langle BEHAVIOUR\rangle :: = [W|S|F|R|Fg|Dk|Fi...]
\]

\[
\langle EGO's ALTER'S ORGASM\rangle :: = [N]|[X][O][H][M][C]
\]

\[
\langle MODIFIER\rangle :: = [null/ \langle associated object list\rangle \langle modifier\rangle]
\]

\[
\langle OBJECT\rangle :: = [P,I,D,T,...]
\]

The symbol \(\langle\) may be read as ‘can be replaced by’ or ‘consists of’; it links the basic term (definiendum) on its left hand side and its specification on the right. The symbol \(\rangle\) may be read as ‘or’ or ‘such that’. The most fundamental units (behaviours, modes, orgasm/ejaculation, conjugators and objects or ‘accompaniments’) are specified as a list of letters denoting the contents in the code. In brief:

A sexual act consists of:

- a behaviour [masturbation, fellatio, ...], where
  - the mode is [self, active, passive, mutual, ...]
  - ego’s ejaculation and alter’s ejaculation can occur in a specified manner
  - in him, on me, in a condom... to be explained later,

  together with

  - modifiers such as [poppers, lubricants, ...].

When the respondent is asked to specify these details (in natural language) for each act he will write, for example:

‘I fucked him; I came in him; he didn’t come.’
which is interpreted as:

behaviour anal intercourse; modality active;
ego’s orgasm in alter; alter’s orgasm no
(modifier/s none).

The specification of the sexual act is open; the list of behaviours can be extended as new activities, such as inter-femoral intercourse, are encountered and such is the imaginativeness of human sexual behaviour that even modality is not closed. 13

The Sexual Code

The structure of the sexual session can readily, simply and efficiently be represented in an encoded form as a string, which greatly simplifies storage and facilitates data analytic operations like comparing the structure of two sessions. The details of the coding system in this context would be rebarbative and need not detain us; suffice it to say that the coded form is isomorphic with the structure of the act, and has an easily remembered mnemonic form. Earlier and simpler versions of the code were used by respondents themselves as an encrypting device and some of these are still used to save unnecessary repetition. 14

Sexual Diary Data

The problems of representation and analysis which arise in the case of sexual diaries are similar to those encountered in content analysis, where one is faced by a large amount of natural language material subdivided in various ways (sources, chapters, sentences and words) each with their own syntax and semantics. The same is true for sexual diary information, although coding is fortunately an easier prospect because the structure of sexual behaviour is (by design) much more tightly defined than ordinary language, and much complexity has been reduced by the preliminary coding.

To illustrate the high degree of structure and redundancy in a typical set of sexual diary data consider this coded version of a week of a respondent’s sexual activity (each session is a sentence ended by a period). 15

| PF, NM.l AS&S PS PRI ARI ATF AF, HN/1 PF/1 HW, NI. (PS AS PS ACP PRI ARI AF, HN/1/HW) | P.RI PF, NM/1 PW, AW PW, AS, PF/1. HW PW, XN/1 HW, NX/1. PF/1. PF, NM/1. | PW PRI PF, NM/1 (PS ACP AF, HN/1/HW) | PW AFG HW, NX AF/1. |

Context and Acts: Data record and string

The original diary version of third sentence/session above was:

[... when we got home after the club I stripped him and, kneeling before me] he sucked me. After moving to the bed, I sucked him. Then I lay back and he started again sucking me. Lying face down I then took the belt to him and began using it on him and after a while he then moved round and began rimming me. I did exactly the same to him, preparing him. Then, with him lying on his back, I then sucked him using KY whilst he wanked himself. I came in him; he didn’t come. We used poppers throughout.

Clearly, the purely sexual information has been abstracted and some incidentals (such as stripping and positional information) have been removed. Each session is a separate data base record 16 as follows:

| NO: cf****/1 | TYPE: vi | STATUS: Neg | DAY: tues | DATE: ***** |
| TIME: 0100 | PLACE: home, after the club |
| PERSON: F2 Regular, 31, sex 4 yrs, Neg |
| ACT: (FS AS PS ACP PRI ARI AF, HN/1/HW) |
| POPPERS: none | CONDOMS: none |
| OTHER: leather belt | DRUGS: alcohol beforehand |

Each field of the record is in upper case (FIELD:) and the entries are in lower-case. Identifying information is denoted by asterisks. The contextual information makes it possible to select out subsets of data with particular characteristics of descriptive or explanatory relevance. Some of the more important ones evident from the above fields of the record include:

- Home area and ID (cf. denotes Cardiff, **** is ID number and /1 refers to Wave 1).
- Age-group and Relationship type of Ego (type vi signifies ‘Over 39 and in an open relationship’).
- Time: day (to see whether sexual activity is different on Wednesdays as opposed to Saturdays); hour (to allow sequencing within a day but also to enable contrast between morning, afternoon, evening and early morning sex).
- Place: Most men are not co-habiting with a sexual partner, so location of sexual activity is interesting. This field also allows for outside sex, for example, parks or toilets to be selected.
- Person: each current sexual partner is allocated a sequential number and
are currently under systematic investigation in Project SIGMA, and any results must therefore be tentative. In these investigations various methods are used. For a balanced subset of SIGMA panel members, comparison is made of the number of sexual acts per month given in the current interviews, and the estimates derived from their sexual diary (these refer to adjacent months, of course). A specially recruited national sample completed a month diary and returned it. After its return they were sent a form asking them to estimate the number of times they had done a set of sexual acts during that month (and given an incentive to do so). Comparison was thus possible for the same month period between a diarist’s counts derived from his sexual diary and the numbers estimated directly by him. Studies were made of the encoding process (both in the cognitive and the technical sense). In the former sense, the focus is upon how people differentially perceive and ‘chunk’ visual or verbal stimuli (of sexual behaviour); in the latter sense the focus is upon how Project coders turn subjects’ accounts into their coded version.

Even at a preliminary stage the results of this research radically undermine the common assumption that data derived from questionnaire/interview data are more reliable and valid than diary data.

Reliability

Reliability refers to stability and internal coherence. Internal reliability in the psychometric sense of alternative forms or split-half reliability are not feasible proposals for diary entries; test-retest methods are virtually impossible to implement and can generate understandable and counter-productive hostility among diarists. Moreover, any attempt to implement independent or repeated entries fast run up against problems of memory recall for detailed data (Linton, 1986) and hence confound any reliability estimation. In SIGMA studies we have therefore concentrated on inter-coder reliability (how far do different coders encode natural language diary entries in the same way) and direct coding (from visual stimuli). Coders are presented first with a spoken or written account from a given diary, and asked to encode it (separately and without collaboration) according to the rules of representation. The coded versions are then compared; inter-coder agreement is assessed by comparing the codings (which are strings) and measuring how similar they are to each other by using a Levenshtein distance (Sankoff and Kruskal, 1983). Distance (dissimilarity) values are generally excellent, averaging better than 0.10.

The visual experiments consist of showing two three-minute sections from a gay porn video collection called Gay Weekend II, produced in the United States but brought into Britain from the Netherlands. Experiments have so far been restricted to project staff and coders. The main purpose is to see how, given a specific visual stimulus of homosexual behaviour (a sequence taken from the video), experienced coders ‘chunk’ this continuous and sometimes ambiguous
material in terms of the Project coding schema (which works in terms of a sequence of discrete events) and go on to ask how similar these accounts are. A major problem that emerges is how to reconcile codings which differ primarily in ‘fine-ness’ or detail, where, for instance, one coder will report that nipple fingering was followed by anal fingering, whilst another will interpose active masturbation between these two acts because the hand went over (and possibly fingered) the penis. It is still too early fully to assess these visual experiments, but acts which resulted in ejaculation are always recorded, and in the correct sequence, which is reassuring.

Validity

Short of observing sexual behaviour directly, it is not immediately obvious what forms validation can or should take. Since SIGMA diarists filled out their diaries in the months following their yearly interview it is possible to identify each diarist’s interview data and compare the two sources. The most cited test treating validity (Campbell and Stanley, 1963) properly distinguishes internal and external validity (or, confounding and generalizability problems). Where there are two or more forms of instrument or method, each with its uniqueness component, they have also stressed the role of triangulation to deal with convergent validity.

Comparing interview and diary estimates of the same behaviour

Several forms of investigating validity are currently being used (see ‘Reliability and Validity’). These centres primarily on the extent of agreement between accounts — derived from different methods (interview estimates vs. derived diary counts) and from different individuals involved in the same sexual session. Internal threats such as instrumentation and selection biases are particularly liable to occur in these contexts, and selection bias in particular undoubtedly occurs among volunteers who agree to do sexual diaries. However, since the SIGMA diarists are a subset of the main panel, demographic and other factors (such as higher rates of sexual activity) which differentiate the two groups can be used to estimate the degree and sources of selection bias.

The main focus in the SIGMA studies is on comparing the similarity between accounts and finding out if (and how) different estimates of the same sexual activity vary. At the aggregate level (i.e., incidence figures over all individuals) agreement is excellent and gives an important clue about the likely relationship between interview estimates diary counts (calculated from the daily-completed diaries; see Coxon and Coxon, 1993b). In particular, rank-order correlation between the aggregate diary counts and the aggregate interview estimates is modest ($r = 0.77$), but the linear correlation is considerably higher ($r = 0.93$). The regression of the diary counts on the interview estimates gives a regression coefficient close to unity ($b = 0.94$) and an intercept of 15.5. Taken together this suggests that, overall, respondents consistently over-estimate the frequency of their sexual activity in their retrospective recall in the interview situation by a constant addition (intercept). The aggregate data are presented in Figure 8.1.

![MONTH INCIDENCE OF SEXUAL BEHAVIOUR](image)

**Sexual Behaviour Codes:**

**MODALITY (1st letter):** Active Passive Mutual/Simultaneous

**BEHAVIOUR (Subsequent letter/s):**

- Code 'S' Street Term: Behaviour
- P Fuck
- S Suck
- W Wash
- R Rim
- T Thigh Fuck
- B Body Rub

**Figure 8.1:** Aggregate diary vs. interview estimates of sexual behaviour

Until recently, the main alternative explanation for any observed differences between interview and diary counts was that the two months were not the same; the interview questions referred to the month past, whilst the diary was not
started until after the interview, and the data hence referred to the subsequent month. In the most recent experiment, however, this has been overcome; on return of their diary, diarists were sent a new (and unannounced) sheet asking them to estimate the number of times in the last month (i.e. of the diary) they had done various acts, and how sure they were of each estimate. In this way the retrospection involved in answering the interview question is simulated, whilst a separate estimate is available by counting frequencies the diary.

Comparing partners' accounts

Another type of convergent validation consists of two participants' accounts of the same event — in this case, the accounts of the sexual partners. Although sexual partners of SIGMA diarists are not identified by name and only some of them are themselves SIGMA diarists, it is sometimes possible to identify partners. Occasionally this occurs naturally where one of the partners mentions the other by name, and he is prepared to solicit his partner's co-operation. But normally this is not so, and the sexual partner can sometimes be inferred by matching attributes (since the attributes of sex partners are described in the diary account) or by matching the contents of the actual sex session itself. This method has been used with a degree of success to estimate 'number of partners of partners' for epidemiological modelling (see Coxon and Joyce, 1993). It is proposed to do a systematic search for such partners and compare the structure and 'line-ness' of the two codings of the same events, and thus gain information not only on the general convergent validity, but also on the detail of what acts are most reliably reported.

An Example: Risk Behaviour in Anal Intercourse

We proceed now to an example of how the diary procedures are used in a specific research problem, namely that of anal intercourse and risk. The primary source of risk in sexual transmission of HIV among gay males is anal intercourse. Safer sex messages advise that risk should be minimized by the use of a condom, in conjunction with a water-based lubricant. If risk is to be monitored it is crucial to ensure that information about involvement in anal intercourse and the use of condoms is accurate, reliable and valid. The usual way of investigating this central issue is by the interview or questionnaire method, with the interviewer defining levels of sexual risk for the respondent and then asking whether or not he has engaged in any of the risky behaviours, and if so, how often.

There are good reasons for expecting that such data will underestimate the actual incidence of risk as respondents will tend to deny or underestimate such infractions of safer sex, and prevalence figures from interview or questionnaire data are best taken as lower estimates. Diary data offer definite advantages for investigating such issues. First, the estimates are obtained by the researcher post facto from the diary accounts, and there is no need to involve the respondent in identifying risk occurrences at all, since these, too, are identified post facto by the researcher. Interviewer and social desirability effects can thus be avoided.

In monitoring the incidence of anal intercourse and use of the condom, most studies (including SIGMA) usually rely upon the respondent to report upon his condom use, and often content themselves with a graded scale of frequency of condom use. But, again, respondents' reports and estimates from interview contexts are liable to be defective and subject to systematic distortion. Indeed, few can reliably give accurate estimates of their condom use, and yet fewer can remember which occasions involved their use.

Characteristics of anal intercourse

As a preparation for the risk analysis let us turn first to some general findings about anal intercourse derived from analysis of the diaries.

- If anal intercourse occurs, it is typically an end-marker to a sexual session. This is especially so when it culminates in ejaculation on the part of one or more of the partners. Once ejaculation has happened it is unlikely that any significant sexual activity follows, and if it does, it often simply establishes a reciprocation.

- Anal intercourse plays a crucial role in determining the power aspect of a session; Davies and Coxon (1990) show that a sexual session tends consistently to be either reciprocal (symmetric) or dominant (asymmetric). In the reciprocal session the pattern of events follows an A, B alternation ('I do it to you, and you then do it to me', or vice versa), in the dominant session power positions are established: one partner becomes the active partner and remains so throughout the session. Fucking is often the Lynchpin of such sessions.

- The pattern of acts which immediately precede anal intercourse has a clear structure: In an earlier paper we have been able to show that the rule is, 'If I fuck a guy, sucking is most likely to precede it (rather than wanking) and I'm most likely to be sucked by him first' (Coxon and Coxon, 1993b). Correlatively, passive (receptive) anal intercourse is typically preceded by fellatio, and usually by active fellatio.

Ejaculation and Risk

Since many sexual acts carry the potential of orgasm, ejaculation is allowed for in the schema and in the coding. Initially, the only distinction noted was whether or not orgasm occurred, since the question of which partner did so was normally
clear from the context of the behaviour and its modality. Next, the outcome suffix was made two-place — the first for ego, the second for alter — making the detail of mutual orgasm easy to code and clearing up residual ambiguity. Finally, the two-place form was further differentiated to specify the actual destination of the ejaculate (semen), so that even potentially risky behaviours (such as alter masturbating himself and ejaculating on ego's body, which might have lesions . . .) could be explicitly identified. To do this we have developed the form shown in Figure 8.2.

<table>
<thead>
<tr>
<th>CODE</th>
<th>Ejaculate Destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>No ejaculation</td>
</tr>
<tr>
<td>X</td>
<td>Elsewhere (e.g. on floor, or unspecified)</td>
</tr>
<tr>
<td>Q</td>
<td>On him (alter), i.e. skin or body surface</td>
</tr>
<tr>
<td>I</td>
<td>On me (ego)</td>
</tr>
<tr>
<td>C</td>
<td>Into a Condom</td>
</tr>
<tr>
<td>H</td>
<td>In him (alter), either anally or orally</td>
</tr>
<tr>
<td>M</td>
<td>In me (ego), either anally or orally</td>
</tr>
</tbody>
</table>

Figure 8.2: Alternatives for destination of ejaculation.

The seven outcomes are arrayed as a weak order of risk of HIV transmission: no ejaculation (1) — ejaculate elsewhere (2) — ejaculate on a partner (3, 4) — ejaculate in a condom (5) — ejaculate in a partner (6, 7). This order differentiates the ‘On rather than In’ safer sex message and the full form also specifies which partner receives the semen. The fifth alternative (the condom) might at first sight seem anomalous, but belongs there as perhaps the most important ‘destination’, being the main form of prophylaxis. These ejaculation (outcome) codes will now be used to examine risk behaviour involving anal intercourse.

‘Volume’ Analysis of Risk and Anal Intercourse

The basic data used for the risk analysis are all the 2107 acts of anal intercourse in the data base and the unit of analysis will be the act rather than the individual (so-called volume or outlet analysis; Coxon and Coxon, 1993b).

Anal intercourse accounts for 5.6 per cent of all sexual acts of gay men in the database (13.4 per cent of all acts excluding masturbation), almost exactly evenly divided between the active/insertive and the passive/insertee modalities. (This percentage markedly differs from the usually reported incidence per man during a month; about a third of gay men engage in either active or passive anal intercourse a median number of two times a month according to SIGMA studies (SIGMA, 1990)).

To investigate levels of risk in anal intercourse we need first to examine the frequency with which the various forms of ejaculation occur. The percentage distribution is given in Table 8.1. The information in this Table reveals several interesting points. One is that coming on a partner is actually a very uncommon activity; ejaculation elsewhere (basically, avoiding the partner's body) is a much more likely eventuality.

<table>
<thead>
<tr>
<th>Code</th>
<th>Outcome</th>
<th>Per cent</th>
<th>Risk level</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>No ejaculation</td>
<td>30.5</td>
<td>0</td>
</tr>
<tr>
<td>X</td>
<td>Elsewhere</td>
<td>12.4</td>
<td>1</td>
</tr>
<tr>
<td>O</td>
<td>On him (alter)</td>
<td>1.2</td>
<td>2</td>
</tr>
<tr>
<td>X</td>
<td>On me (ego)</td>
<td>1.7</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>Into a Condom</td>
<td>16.6</td>
<td>4</td>
</tr>
<tr>
<td>H</td>
<td>In Him</td>
<td>23.5</td>
<td>4</td>
</tr>
<tr>
<td>M</td>
<td>In Me</td>
<td>13.4</td>
<td>4</td>
</tr>
</tbody>
</table>

(N Actuals = 2107)

Table 8.1: Destination of ejaculate in anal intercourse.

Putting the information hierarchically — out of 100 acts of anal intercourse, 70 involve ejaculation of semen. Of these ejaculations: 17 are into a condom, 53 are unprotected, 12 go elsewhere, 4 go on the partner's body and 37 go into the partner's anus. The actual level of high-risk behaviour revealed is disturbingly high since over one-third of the acts of anal intercourse are in the highest risk category of all. High-risk sexual behaviour thus appears in this analysis to be much more prevalent than normally reported in studies of gay men, possibly because high-risk behaviour is concentrated in particular individuals. In the next section we shall therefore enquire whether there is significant variation among particular types of gay man.

Condom use also presents a disturbing picture. It is well-known that condom use is far from universal among gay men who practice anal intercourse. But the figures in Table 8.1 derived here from their diary accounts shows that if ejaculation occurs in anal intercourse, a condom is used in less than a quarter of the instances. Moreover, protected ejaculation (into a condom) occurs only half as frequently as the most high-risk behaviour of unprotected ejaculation.

Have these fractions changed over the four annual waves represented here? Table 8.2 presents the relevant information. The percentage of protected acts of

<table>
<thead>
<tr>
<th>Anal Intercourse</th>
<th>WAVE 1</th>
<th>WAVE 2</th>
<th>WAVE 3</th>
<th>WAVE 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condom [C]</td>
<td>20</td>
<td>13</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>Unprotected [H,M]</td>
<td>33</td>
<td>33</td>
<td>36</td>
<td>37</td>
</tr>
<tr>
<td>Ratio: [C]/[H,M]</td>
<td>0.61</td>
<td>0.39</td>
<td>0.50</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Table 8.2: Percentage of acts where anal intercourse involves a condom [C] or is unprotected [H and M] by wave.
anal intercourse (i.e., where a condom is used) actually drops by a third in the second wave (a phenomenon noticed elsewhere), and only climbs back up again over the next two; the fraction of unprotected acts of anal intercourse actually increases over this period. The ratio of protected to unprotected acts was highest in the initial wave (1987), fell dramatically in the second wave and although increasing subsequently has not yet reached the initial level.

The Effect of Age and Relationship on the Volume of High-Risk Activity

Having found this variation, how is it to be explained? Power and sex-role almost certainly have a part to play (Coxon and Coxon, 1993a), and knowing that receptive anal intercourse is the more risky would normally lead us to differentiating active and passive modalities of anal intercourse in the analysis before going any further. But instead, in the space available, it may be more profitable to see how unprotected anal intercourse varies by the two factors that we have found to be repeatedly the most potent in explaining variation in gay men’s sexual behaviour — age and relationship-type (Coxon, 1987). Because these factors are themselves associated, it will be best to take their combined effect on the unprotected anal intercourse.

The nine-fold SIGMA typology of three age-groups by three relationship types is now used to look at the percentage of acts of unprotected anal intercourse within the nine cells. This is presented graphically in Figure 8.2 with Table 8.3.

<table>
<thead>
<tr>
<th>Relationship Age</th>
<th>Less than 21</th>
<th>21-39</th>
<th>Over 39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closed Relation</td>
<td>30.1</td>
<td>34.7</td>
<td>55.1</td>
</tr>
<tr>
<td>Open</td>
<td>27.8</td>
<td>24.5</td>
<td>38.2</td>
</tr>
<tr>
<td>No Regular</td>
<td>0.5</td>
<td>22.0</td>
<td>40.0</td>
</tr>
</tbody>
</table>

Table 8.3: Percentage of unprotected acts of anal intercourse, by age and relationship type.

There is clearly a very substantial amount of variation in these rates of unprotected anal intercourse. High-risk activity is positively related to age (the older groups have increasingly higher rates). Relationship also has a systematic effect: highest-risk anal intercourse is highest in closed relationships, slightly lower in open relationships, but considerably lower among those who have no regular relationship. This conforms well to findings from the panel-based data of the project (Weatherburn et al., 1991; Hunt et al., 1993), where issues of intimacy and trust rather than recidivism or relapse (Davies, 1992, 1993) are shown to explain such differences. Nonetheless, it is important to have some idea of whether age or relationship is the greater influence, and where the most important combinations occur. From the start of diary analysis (Coxon, 1987) we have used a resistant method of analysis to answer this question; means analysis with a median polish (Mosteller and Tukey, 1977: 165–202).

The analysis is a simple additive model: the table entries are the sum of the overall average (typical value TYP, here 27.8 per cent) plus the row (relationship) effect, plus the column (age-group) effect plus the combination/interaction/residual effect. In terms of main effects, age-group has a greater effect than relationship type, but this is primarily due to the older group (having an effect almost two-thirds the size of the overall effect). Being in closed relationships have the highest and positive effect. The most striking thing about the combination-effects is that the youngest age group who are not in a regular relationship have a remarkably lower rate — again, almost two-thirds the size of the overall effect. In brief, there are major variations in the rates of unprotected sex among these types of gay men; being older men (40 and over) and (to a lesser extent) in a closed relationship increases the rate of high-risk sex, and having no regular relationship and being young (under 21) serves to decrease the rate. Over and above this, young men with no regular relationship have a considerably lower rate than would be expected.

<table>
<thead>
<tr>
<th>Relationship Age</th>
<th>Under 21</th>
<th>21-39</th>
<th>Over 39</th>
<th>Row Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closed Relation</td>
<td>7.9</td>
<td>0</td>
<td>-4.3</td>
<td>6.9</td>
</tr>
<tr>
<td>Open</td>
<td>-16.0</td>
<td>0</td>
<td>4.6</td>
<td>-5.8</td>
</tr>
<tr>
<td>No Regular</td>
<td>4.6</td>
<td>0</td>
<td>18</td>
<td>TYP - 27.8</td>
</tr>
</tbody>
</table>

Table 8.4: Median polish of percentage of unprotected acts of anal intercourse, by age and relationship.
On this analysis of the sexual diary data, then, the youngest age group who are in no regular relationship have a very markedly lower rate of unprotected high-risk sex than any other. This is an encouraging (if not entirely unproblematic) finding, and accords with other trends documented by SIGMA (Davies et al, 1992): young gay men conform far more to safer sex guidelines than any other group.

Conclusions

The diary method provides an excellent tool for the study of sexual behaviour, yielding fascinating detailed and often unique information. If combined with an explicit theory or schema of sexual behaviour, comparison with other methods is straightforward. However, it is undoubtedly a demanding method (on both respondents and analysts) even if analysis is aided by recently available software. Validity studies suggest that while aggregate data correlate extremely well with those derived from other methods there is an undoubted selection bias in those who agree to keep such diaries.

The most striking substantive conclusion is that there is a remarkable amount of highest-risk activity (unprotected anal intercourse to ejaculation) among gay men — far more than would be suspected from individual-based analyses. This highest-risk activity is concentrated in older gay men and is markedly under-represented in young gay men who have no regular relationship. It would have been impossible or totally unfeasible to use questionnaire or interview methods to obtain this striking result.

Notes

1 This rather ugly clause is more accurate than homosexual or gay, which imply restriction to this orientation, or identification with it. In the rest of the chapter gay will be used to denote behaviourally homosexual males.

2 The presence of a third person (here the social scientist) would render homosexual activity illegal under the 1967 Sexual Offences Act as being no longer a private activity. Scottish law does not contain this restriction.

3 Project SIGMA (Socio-sexual Investigations of Gay Men and Aids). The research in this article was funded by the Medical Research Council and by the Department of Health (UK). Work on Sexual Diary software and on the validity and reliability of the sexual diary method reported here is funded under a series of grants from the Department of Health. The views expressed in this article are those of the author and not necessarily those of the Council or of the Department.

4 We have been able to show that it is selection bias toward those with busier sex lives rather than exaggeration of activity that is primarily producing bias.

5 The male pronoun is used descriptively throughout this chapter as we are referring exclusively to gay men.

6 As a rule of thumb, this is usually less than a three-day period for detailed accounts of sexually active men. Particular, care or risky behaviour can often be recalled for a longer time, but even here the contextual information and surrounding sexual activity is poorly recalled (see Brewer, 1988; Neisser and Winograd, 1988).

7 The ethical undertaking given by the Project was that the name of someone identified in an interview would not be part of a Project record. This was undoubtedly important in maximizing response and trust initially, but prevented any record linkage and severely limited network analysis. The ethical issue is discussed in Coxon et al., 1993c, and the practical network issues in Coxon and Joyce, 1993.

8 But also because they could have such an effect: condoms can act as a barrier (and are anyway encoded into the outcome information), lubricants can increase the probability if not water-based and may decrease it if containing Nunyung-9; toys such as dildoes, if shared may potentially transmit HIV virus.

9 This distinction is in common use in the gay community. It is not to be confused with the intercourse distinction often used in this context. Many (indeed most) sexual acts are not penetrative.


11 Orgasm and Ejaculation are used interchangeably, though it is recognized that they are not necessarily equivalent.

12 The entries in square brackets are instances taken from the appropriate list. The letters in the behaviour list are the first letter of the street term of the behaviour, e.g., "P" for "Pawking, Sucking..."

13 We had to include a new modality ‘H’ to deal with situations where alter did something sexual by himself, which could still have relevance to HIV transmission. An example is where he masturbates himself but ejaculates on his partner.

14 The most common activity is solitary masturbation to ejaculation/orgasm. This is now well-known in the gay community by its SIGMA code of "SWO."

15 Everything but the sexual session data has been removed — time, place, partner, etc.

16 Using either dBase or CARDIOX-Plus software. Selection using boolean operators is done within the database and then exported as ASCII or Basic (comma-delimited) format to specially designed or commercial software.

17 In order to derive information on partner mixing for the Imperial College group of epidemiological modellers (Anderson, 1986) an estimate of the number of partners of partners was needed. The network procedures used to do this on a Diary dataset are described in Coxon and Joyce, 1993.

18 The Programs and data are available at cost from the author.

19 This sub-project on reliability and validity is funded by the Department of Health.

20 The incentive consisted of donating £2 to the Terrence Higgins Trust on their behalf on return of a completed estimate sheet.

21 The same rules are presented in the instructions to diary keepers (SIGMA, 1993). Originally respondents were invited to encode some or all of their data but this is now discouraged except for very common events such as solo masturbation.

22 SIGMA uses the 7-point gradation: Always / Almost always / More than half the time / Half the time / Less than half the time / Very seldom / Never.

23 The direction of distortion can be upwards or downwards, though it is more likely to be downwards. Upward distortion is a kind of Don Juan effect: boasting about one’s sexual prowess, and downward distortion usually arises because anal intercourse is known to be most risky and subject to disapproval by health education authorities.

24 It is important to stress that this is a contextual and not an individual property. Some individuals will always engage in one of the two types of session or always adopt an active or passive role, but these are in a minority. Most gay men shift modality in different settings or with different partners or even within the same session.
25 Throughout this and following sections of the chapter the phrase 'unprotected anal intercourse' means unprotected anal intercourse in coitus.

26 This group is relatively small in size (31/569); spread across 4 waves it is difficult to see a stable trend. Selection bias could account for some of the unexpected very low values, but not the relative amount.

References


SIGMA (1993) 'Notes on keeping a sexual diary', Colchester, Project SIGMA.

Networks and Sex

The Use of Social Networks as Method and Substance in Researching Gay Men's Response to HIV/AIDS

Anthony P.M. Coxon

This chapter is based upon the view that much behavioral research on gay men and their reaction to the AIDS pandemic is not so much psychologistic as overly individualistic in approach and that this severely limits its utility and prevents us from tackling the most challenging substantive and methodological (and hence policy-relevant) issues.¹

This bias in the design of such studies often begins with the words: "Since a random sample of gay men is not possible, we used snowball sampling..."

They usually didn't— and the bias continues into the heartland of sexual behavior. Despite the building up of wave upon wave of prevalence figures for this and that sexual act, we still know little more than Kinsey did about the context of relationships.

A critical deficiency is the lack of a network perspective— network concepts, notions and methodologies. Now it may well be that networks are believed by the research community to be too complex to study and that they do not really add value sufficiently to compensate for their undoubted problems. In part this is true; it is often difficult enough to persuade an individual to participate in sexual behavior studies, without having to go on to recruit couples or— more relevantly—casual partners. It may also be that sex researchers shy away from the greater involvement and commitment of such detailed research methods, and prefer to rely upon the tried-and-tested survey.

I want to explore the applicability of network notions in three areas that loom large in our own research and in many other projects on sexual behavior. Network notions impinge most directly in the areas of:

* design,
* in networks of sexual contacts, and
* in identifying (unknown) sexual partners.
In so doing I shall be frank and realistic, so that others may learn from our mistakes and be a little more adventurous than funding authorities might sometimes like us to be!

1. DEFINITIONS AND SAMPLE SELECTION
From the outset, Project SIGMA resisted any attempt to produce a random sample of gay and bisexual men. This was done on both principled and practical grounds. In principle, the tendency for policy-makers and others to think of "homosexuality" as a lasting and recognizable attribute was not to be encouraged, and Kinsey's research (1948:650–657) illustrates well how prevalence estimates of "homosexual men" can be made to range from four percent to almost fifty percent by successively relaxing the criteria of the type of sexual contact and the time-period of sexual involvement with those of the same sex (Coxon 1987).

But practical issues were paramount; no general population survey of sexual behavior was then envisaged in the U.K. and the cost of attempting to sample randomly on a two-stage basis (initially "combing") to produce a population frame, and secondly sampling within it would be well outside funding agencies' means. In any event, the relevant population is narrower than the notional category of "homosexual men," being concerned with potential or actual HIV transmission rather than sexual identity. By this time, the causative agent of AIDS was known to be viral and the first crucial network-contact study (Auerbach 1984) had been published.

The strategy finally adopted by SIGMA was therefore to structure respondent selection round the two factors known to maximize variation in homosexual behavior:

- Age,
- Type of Relationship.

Enter the network, via Rapoport (1953, 1957), Rapoport and Horvath (1961), Fararo and Sunshine (1964)—and snowballing.

1.1. Sampling Via a Large Social Network: Theoretical Ideas
The conceptual basis of the sampling strategy was motivated by the question of how to obtain systematic (ultimately unbiased) information about a large, unknown, and connected social network of sexually active gay and bisexual men. I was familiar through earlier sociometric interests with Rapoport's models of diffusion through large networks. His original application in mathematical biophysics had been to the neural net, estimating the "gross statistical properties" (Rapoport 1963: 512) of a huge network of unknown size and structure, and in particular the "close-knitiveness" and the "ultimate connectivity" of the neural net. This approach was applied in turn to communication and (by Fararo) to friendships among young and institutionalized offenders.

The aim in these social studies was to explain diffusion within these empirical social networks, the manifest divergence of these processes from diffusion in a purely random baseline network, and to do so by identifying and estimating significant bias parameters (Rapoport 1957, Fararo and Sunshine 1964, Skowron 1985). These parameters, when incorporated into the "reduced axone density" (in effect: the "slowed-down" effective out-degree, referred to as $\alpha$), explained the (lower) ultimate connectivity and rate of increase in new contacts by means of such modifications of a random Polya model.

In this model the expected fraction contacted at time $i$ is given by the iterative equation (Rapoport 1961:285):

$$p_{i+1} = (1 - X_i) (1 - e^{-\alpha})$$

where $p_i$ denotes the expected proportion of new contacts at step $i$ in a random net, and $X_i$ the cumulative proportion contacted at time $i$. The expected ultimate connectivity $X_i$, satisfies the transcendental equation:

$$X_i = 1 - (1 - p_i) e^{-\alpha}$$

where $X_i$ is the asymptotic value of the cumulative fraction contacted. The epidemiological parallel is of course obvious, and in later work Rapoport actually referred to the model as the (Polya) "contagion process model" (Rapoport 1979).

The method for making the estimates of diffusion or "infection" is the "tracing," which consists of producing a rooted tree giving the new contacts at each ordinal step $i$. Starting at a small initial fraction of nodes, the tree is "grown" by consulting the sociometric contact matrix at each new contact, until no new contacts occur. A small illustrative example is useful.

1.1.0. The Tracing
Let us suppose that the network structure is known (which is often not the case, but simplifies the exposition), and is given by the adjacency matrix $A$, whose element $a_{ij}$ is 1 if there is a relational link between points $i$ and $j$ and is 0 otherwise. An example is given in Figure One(A). Here the number of points, $N_i$ is 10 and the out-degree of each point (number of contacts made) is a constant 2. The starting set for each tracing is always small compared to $N$, and in the two examples shown it consists of a single point: point E in Figure One(C)(a) and point E in Figure One(C)(b).
Each tracing proceeds by defining the starting set (step $t = 0$) and then noting the contacts which this starting set makes. These constitute the set of (two) new points contacted at step $t = 1$, and these are counted into $n(1)$ (the number of new contacts at step $t = 1$). The contacts of each of these new contacts in turn are now identified, but only if they have not been contacted previously do they count in $n(2)$. This tracing process continues until there are no new contacts; this stage of “ultimate connectivity” may be complete (contacting the entire set of 10 points as in tracing (ii)) or partial (as in tracing (iii) in Figure One (C)). Notice that the number of points newly contacted at each step and the number of steps taken to reach the end of the tracing differs depending on the starting set.

To emphasize the fact that it is new contacts only that are considered at each step, tracing (ii) starting with point E is represented in tracing (iii) by drawing in the redundant (already contacted) links at each step as a dotted arrow. Because they have already been contacted, they of course point backwards.

A given tracing is characterized by the fraction of the population finally contacted (average ultimate connectivity) and by the number of steps taken to reach this stage. In Rapoport’s theory, a number of tracings are made, until the average tracing distribution (i.e., averaged at each step over all the tracings) stabilizes to within a desired limit. This asymptotic tracing distribution is given in Figure One (B). The proportion of new points contacted at each step is (simply $n(t)/N$) is given in $p(t)$, and the cumulative proportion of points contacted by step $t$ is given in $X(t)$. The “ultimate connectivity” of the network is given by the last entry. Thus, on average, eighty-four percent of the points are ultimately contacted by the sixth step when a tracing is made in this network.

The interesting difference between the neurological and the normal social science applications is that tracings are usually produced in the social
1.1.1. Sociological Implementations

The conceptual analogy with sampling gay men seemed well-nigh complete. Implementing a tracing (or several) provides an excellent methodological specification of what sampling a hidden population should be. If continued to completion (and without error) such tracings would provide an enumeration (of at least connected subsets) of the homosexual population and also information about its local network characteristics. This, of course, an ideal type and is practically unrealistic as a technique in toto. Nonetheless, it tells us what “snowball sampling” should be. A close parallel to the tracing process was known to sociologists from Coleman’s study of diffusion (Coleman 1958), where physicians were asked to name those of their colleagues to whom they gave information about a new drug, a method now known as the “chain-referral” method (e.g. Biernacki and Waldorf 1981).

But there are important differences. In the sociological examples, attention has shifted from Rapoport’s interest in the inherent properties of the network and the mechanism generating it to the usefulness of a strategy for reaching hidden or hard-to-reach populations. To be sure, the underlying assumption is that there exists a network of contacts (for otherwise why would snowballing work?), but even considerations like the number of stages/steps and the outdegree (constant or average number of contacts) slips out of sight.

Paradoxically, most statistical work on snowball sampling (Goodman 1961, Holland and Leinhardt 1979, van Metter 1990, Snijders 1992) has not generally been directly relevant since it typically refers not to populations of unknown size, but rather to issues of inferring population network properties from samples. Few indeed are the studies claiming snowball status which come in any way near satisfying requirements such as sampling to exhaustion. Perhaps, given the wider meaning nowadays given to snowball sampling—in effect, the cumulative but haphazard acquisition of a quota or convenience sample—it is better to use the term “tracing sample” to refer to tracing procedures which follow contacts-of-contacts to the exhaustion of new contacts.

1.1.2. Sampling Gay Men: SIGMA’s Sampling Procedure

The sampling process used in SIGMA was two-stage: first to obtain easily-accessible respondents in each of the nine Project Design typology cells (chiefly from gay pubs, clubs and voluntary organizations); secondly to use these initial contacts as starting samples for producing tracing trees. The imagery sometimes used in SIGMA to refer to this process was “snowing into the iceberg”; indeed, given the fact that those most “out” are a highly biased group representing only the tip of the iceberg, the idea of obtaining less “out” contacts of the same Age x Relationship-type by snowballing is attractive. But notice the looseness of the relational definition we used; in practice the interviewer asked the initial respondents to name other potential respondents who were of the same (Project) type as themselves, but preferably less “out” as gay. It was left to the interviewer to satisfy him/herself that this definition was understood by the respondent, and we were rarely able to ascertain whether this had actually been done.

The attempt by SIGMA to implement tracing sampling was noble, but ultimately deficient, and for a number of instructive reasons:

- Often a given gay man’s friends and acquaintances are not of the same Age-Relationship type as himself, so that it was frequently quite difficult for a respondent to name someone of the same type, let alone someone that was less “out.”
- The number of contacts to be named was never specified; more relevantly, there was no criterion provided by which the respondent could decide when the number of his nominees was sufficient.
- As a Project we had bound ourselves to anonymity in the form of not recording or making use of the name of anyone named in the research context. We therefore had to rely upon the respondent to contact his nominee and ask him to participate in the Project. Consequently we might never know that a specific person had been thus nominated, let alone whose nominee he was.

However, in terms of the stated objectives—to “snowball” into the more covert gay population—a name was some degree of success. The first “Question Schedule” contained a number of questions asking who knew that the respondent was gay/bisexual. Inter alia this provided a useful indicator of “outness,” and (at least in the South Wales etc) this index of “outness” decreased as known contacts were interviewed.

But it must be said that the exercise was not overall a resounding success, and that to all intents the initial SIGMA sample was no more (nor indeed any less) a “snowball” than other such studies. The reasons cited above are enough to account for its lack of success, but in principle each could be
remedied. In Section Three, I present information about how critical the actual naming (identification) process is. But the main shortcoming was that the criterion/relationship for respondent naming was not only too vague, it was also not related directly enough to the sexual transmission methods we were studying. This raises the question of whether a genuine (sexual) tracing sampling technique could have been devised and implemented, and whether it would have been more relevant. Without doubt the attempt might bring to light yet more compelling shortcomings (not least those connected with confidentiality and compliance) and would probably be more expensive in terms of time and money. I shall return to this in the next section; it remains to assess the depth of the burrowing.

The gay scene in Cardiff and area is a good deal more closely-knit (on any significant criterion) than that of the sister site of London. In Rapoport/Fararo terms, this implies higher values of reciprocity ("sibling bias, 8") and transitivity ("parent bias", p). This in turn implies longer chains because the clustering leads to more redundant (already contacted) new contacts at any step, and thus "slow down" the growth of X(t) and of the eventual connectivity asymptote.

Where it was possible to track the contacting process in Cardiff, these redundant new nominations occurred with sufficient frequency that a rule was drawn up that new nominations had to be checked by the site office before being allowed as a new sample member; in London this rarely happened.

For sexual contacts (of whatever variety) there turned out to be a goodly number of cross-cutting circles, but with weak links between them, so that an estimate of ultimate connectivity probably depends rather importantly on whether the sample includes the liaison persons (bridges) that mediate such clusters. Not including bridges would lead to a falsely low estimate of ultimate connectivity (and hence of prevalence). It would also lead to missing certain important subsets of respondents who come in and out of the scene on an occasional basis and who would only normally be contacted via one man; occasional (but not hardened) users of "cottages" are an important example, and a relevant important sub-group epidemiologically.

These assessments of the topology of the homosexual network are largely impressionistic, and would need to be investigated directly as hypotheses. In neither site, however, did we normally exceed a chain-length (let alone a tracing step-length) of more than three, and we argued that a length of four would be necessary to achieve even reasonable coverage, and ten or more would be necessary to come anywhere near encompassing a coherent cluster (Davies 1986).

The question of whether engaging in tracing samples of actual sexual contact (and especially of implicated behaviors such as anal intercourse) would materially improve our network-knowledge and assist in estimating prevalence and prediction is a good deal more moot, and brings us nicely to the second section.

2. NETWORKS OF TRANSMISSION
The main implicated behavior for the sexual transmission of HIV among gay men is, of course, and intercourse, with the remoter (and contested) possibility of fellatio (but see Koopman et al. 1992). An important framework for interpreting and predicting transmission is therefore the sexually defined network. The defining relation can be sexually multiplex (most gay men don't just have anal intercourse) and will normally need to be restricted to a fixed incidence time-period. The sexual relation is also necessarily asymmetric, since the probability of transmission is greater for the passive partner (Darrow et al. 1983, inter alia). This asymmetry means that sexual networks and tracing trees have to rely on directed graphs. Moreover, sexual role, though normally considered as an individual or point property (see Coxon et al. 1993) can more appropriately be viewed as a dyadic property. Thus, although in individual terms a man may engage in only Active, only Passive, Both (or Neither) modality (the so-called BAPN roles), this is often distributed associatively so that a man is (say) only active with one partner and only passive with another (this will still lead to his being classified as "both" in the individual-based role categorization).

It is important to specify why the Rapoport model is being used as a basis. After all, HIV is not generally highly infective, and one contact is rarely sufficient for infection. We also know that epidemiologically it is necessary to take into account the stage of HIV infection in assigning probabilities of infection. What the Rapoport models do is to provide us with a potential infection model—a network structure of routes and paths along which infection can travel, and where (as with a block-model) the zeroes are often as significant as the ones. Many existing epidemiological models are long on necessary conditions for infection but short indeed on the socio-sexual structure which in large part constrains and contains the infection. Equally, these models often require demands for data that normally are not collected, particularly by survey analysis—such as information on mixing (Anderson et al. 1986) and especially on the second-step information on the number of sexual partners of one's sexual partners. These become crucial parameters in the Anderson model.

2.1. Piloting a Study of a Two-Zone Network of Anal Intercourse
At an early stage in the pilot work of SIGMA in 1984 (reported in Coxon
1986) I began investigating the possibility and viability of mounting a network study of anal intercourse within the South Wales SIGMA site. This would be akin to contact-tracing of STD Clinics but more sensitive and sociologically useful. It is relatively straightforward to obtain information about whether or not a gay man engages in anal intercourse, the rates and number of partners in a given incidence period, and the extent to which condoms are used. These are now routine questions in any socio-epidemiological enquiry. Only rarely, however, are the pieces put together so that, for instance, we know whether anal intercourse is engaged in with a particular partner, and with what modalities.

In Project SIGMA we have developed the Sexual Diary as a supplementary method for obtaining precisely this information (Coxon 1988; Coxon et al. 1992, 1993; Davies 1990). The main shortcoming is that by the self-denying ordinance mentioned above, we specifically have not asked for the names of partners (though we do ask for their descriptive attributes), thereby apparently foregoing the possibility of linking the data to obtain sexual networks.12

The other strategy was to attempt to construct an anal intercourse network directly, by following a proper tracing procedure, initially with a single root sample. In this example I refer to one such network, but restrict attention to the initial one-step partners of Respondent #One, although the tracing actually continued beyond this point. The tracing procedure was as follows:

Sex Tracing

- Specify the incidence period (and bounds)
- Choose the root node/s
- Establish how many partners he had engaged in anal intercourse, get their name, information about them, details of act/s of anal intercourse, if possible, including modality, ejaculation and condom use
- For each named partner, repeat step 2.

(The sociogram of the network centered on Respondent One is given in Figure Two.)

Step 0 This is straightforward. In the example it was taken as a six month period, and limited to the South Wales Project area.

Step 1 This starting node is chosen for good (cautionary) reasons. His position is discussed below.

Step 2 It is fairly unusual for a gay man to have thirteen partners with whom he engages in anal intercourse, but by no means rare. It was not difficult to establish information about the initial thirteen linked partners from him, and his account could be checked against his Sexual Diary account. At the level of whether or not he had had anal intercourse, and in what modality, the two accounts concurred.

Step 3 Knowing the identity of the partners, it was possible to match ten of the partners with SIGMA respondents (all except Respondents Twelve, Fourteen, and Four), and it would have been possible to find the inter-zone One links from Diaries or Interviews and not have to rely on contacting them directly. Information could only be obtained for the three non-SIGMA respondents by direct contact, and in one case at least (Twelve) we knew that he had already refused to be a part of the Project panel, and hence was highly unlikely to give information on his anal intercourse partners. Beyond zone-One (i.e. three-step and beyond) I have only indicated contacts outside the initial thirteen.

Several comments are in order:

- The root node (1) is a professional man in his mid-thirties, with two regular partners (two and three); six, eight and ten are more-or-less regular partners—"occasionals" might be a better term. Partner Five is originally an "affair" (regular partner) of Two's. The rest are mostly casual partners, but include a few more-than-once casuals.
- In terms of individual sex-role, almost all of the men are "B" [both active and passive] during this period, but Six is A[active only] and Nine is P[assive only].13
- In terms of the sociometric structure,11 the core "clique" (maximally
connected symmetric subgraph) consists of \{1,2,10\}, with \{3\} as an important appendage. Partner Nine is a particularly interesting (and at-risk) person, receiving by far the highest number of direct and indirect paths of anal intercourse, whilst Ten by contrast generates the maximum number of anal intercourse paths (see Appendix 1).

- In terms of the Rapoport bias parameters, reciprocity is 0.17 (fairly close to the proportion of “Both” in larger studies [Coxon et al. 1993]), and transitivity (“parent bias”) is 0.04, which is naturally attenuated in this restricted set, but which is not much less than the value for larger networks—after all, anal intercourse is only likely to be (trivially) transitive among those who are both active and passive (e.g., \{1,2,10\}), though “genuine” (transitive in one direction) chains do appear, as in \{1,7,8\}).

- In 1985 (to which the data refer) there is no instance at all of a condom being used in this group during the six-month period, but the HIV prevalence rate was also very low in this non-Clinic South Wales subsample (Hunt et al. 1990).

- This small world is by no means self-contained. At the second, and especially at the third step, the contacts begin to leave South Wales, and also include female partners. Partners \{10, 3, 13\} have a female partner, and \{5, 3, 7, 6\} have one or more male partners in London, which has always had a higher HIV sero-prevalence rate than South Wales.

It is very dangerous to over-interpret these small and illustrative data; the point of the exercise has been to show that it is a potentially valuable exercise, albeit a difficult one. In particular, it involves considerable data-collection costs and requires extraordinarily sensitive and delicate skills on the part of the interviewer.

With the adoption of the self-denying ordinance in not recording or using names, the attempt to implement anal intercourse tracing was virtually abandoned, not least because the funding agency disallowed it—on financial, not ethical grounds, it should be said.\(^{15}\) So: was this another resounding non-success? Let us see.

2.2. Inferring Partner Identity

Perhaps the most galling thing about the ordinance about anonymity was that in many instances, respondents \did\ give names of partners. The reluctance was not because they did not want to name partners (at least under conditions of confidentiality), so much as their concern that the nominee would find out that he had been named by the respondent. But even when the respondent did not name his partner/s it was sometimes possible to infer his (or her) identity from circumstantial or public knowledge—interviewers themselves were usually involved in the local gay scene—and even from within the interview. The accuracy of such inferences and guesses was much aided by the fact that although names were not asked, descriptive information was.

2.2.1. Partner Characteristics: Matching Attributes

From the Wave One “Question Schedule” (1987) onwards, respondents were asked to give a range of attributes for their regular partner/s—Sex, Age, Race, Job, Marital Status and Domestic—and throughout the Project all Sexual Diary keepers are asked to describe each partner during the month in terms of the following characteristics:\(^{16}\)

<table>
<thead>
<tr>
<th>Partner Attributes (Sexual Diaries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 Partner (sequential number and Sex)</td>
</tr>
<tr>
<td>1 Status (Regular, Occasional, Casual/One-off)</td>
</tr>
<tr>
<td>2 Age (known or guessed)</td>
</tr>
<tr>
<td>3 How long you’ve been having sex</td>
</tr>
<tr>
<td>4 Where met (on this occasion) [Casual partners only]</td>
</tr>
<tr>
<td>5 Other information—Job, basis of attraction, payment?</td>
</tr>
<tr>
<td>6 HIV status, if known</td>
</tr>
</tbody>
</table>

Members of Professor Roy Anderson’s team at [then] Imperial College had expressed interest in using the Sexual Diary data for making estimates of mixing ratios—i.e., of information not only about the number of partners, but also of the number of partners of partners.\(^{17}\) If partners were identified by name (and, more demandingly, if they were Project diary-keepers themselves), then such a procedure of estimation depends simply on establishing linkage between a respondent’s file and those of his partners. But for reasons explained in the last section, the SIGMA Undertaking on Confidentiality meant that this was not possible. Dr. Chris Joyce originally suggested a possible strategy: to use information which a diarist gave about a given partner as a yardstick or template “profile” and then attempt to identify him by searching for (preferably one) respondent whose data matched that information.

One option (not systematically followed out) was simply to use the partner attributes, so that if, for example, Jim had described his Partner Five as:

> A regular partner / aged 24 / with whom I’ve been having sex for 2 years / who is a well-endowed teacher from Manchester / and HIV antibody negative

then a SIGMA diarist, Fred, who was found to have the same characteristics would be identified as putative Partner Five, and the information he gave about the number of his sexual partners would then be derivable from his Diary record. The main difficulties are:
Social and Sexual Networks

- What counts as “the same characteristics,” especially given that the fourth category (“Other information”) is open-ended?
- Will a pair describe themselves in the same way? (Thus, Jim’s “regular” partner Fred might consider himself an “occasional” partner of Jim.)
- How much “noise” / tolerance for error can be allowed in matching, especially given the fact that reported age is often systematically biased downwards?
- What happens if more than one candidate appears? Suppose Fred1 is “Regular / 28 / sex for 2 years / student from Stockport / HIV neg.” and Fred2 is “Occasional / 26 / sex for 2 years / student at Salford University / HIV neg”; which shall be identified as the “real” Fred?
- In the case of casual or one-off partners, it is quite likely that information will be very deficient and grossly misperceived, making matching virtually impossible. And yet these cases may often be very important epidemiologically.

Matching by attributes (and especially on open-ended and multi-reference ones) can therefore be a very hazardous procedure, and if there are too many mismatches then the subsequent constructed network will become almost misleading and very possibly useless.

But all was not lost; rather than tighten up or extend the partner characteristics, a different tack was employed.

2.2.2.2 Retrospective Networking

The next ploy was to ignore such individual characteristics (at least initially) and concentrate on much more detailed, relevant and less-easily matchable data. This consisted of “retrospective networking”—namely matching a partner by matching the sexual session itself in which they had both participated after the event, an obvious choice to make!).

The process had three steps: screening, temporal matching, and behavioral matching.

2.2.2.1 Screening

First, all solitary sessions (involving only one partner, i.e. the diarist) were removed, which amount to exactly one-third of the sessions: 33.3 percent (Coxon 1990:15) in this set of data.

2.2.2.2 Temporal Matching

Obviously, a match should only occur if the sessions involved had taken place at the same date and time. However, some leeway had to be allowed on this, since earlier studies had shown that displacements up to one day between the partners’ accounts were not unusual. The matches produced by this process were then treated as fulfilling a necessary condition for matching.

The results of purely temporal matches are presented in Figure Three.

![Figure Three: Number of Temporally Matched Sessions and People Over Diaries Epoch](image)

- First, all sexual sessions involving only one person (e.g., solitary masturbation) were excluded, leaving 2783 sessions for 166 diarists.
- These sexual sessions occurred on a total of 359 days, with an average of 7.6 per day (and a maximum of 66).
- The maximum number of diarists reporting one or more session in any one day is forty, with an average of 5.4 diarists/day.
- A possible match consists of any pair of sessions which occurred on the same day and which could therefore represent the same session written from each partner’s viewpoint...
- ... but this means a large number of potential matches: for these data, either 15,938 matched sessions, or 14,374 potential partners; a truly impressive, but unrealistically large, number.

2.2.2.3. Behavioral Matching

Having established a potential match by temporal matching, the next stage consisted of matching a given session and a potential identically referring (matching) session by comparing the strings (encoded sequence of sex acts) embodying that session.

Before explaining this process, a brief excursus on the encoding of sexual behavior is in order (see Coxon et al. 1992).

2.2.2.3.1 Encoding the Sexual Session as a String

Diaries are initially written in everyday language according to a specified
structure (see above). On receipt they are then encoded into a more efficient form as a quasi-linguistic string (Coxon et al. 1992). An example follows:

Diarist’s Account:

We’d been drinking in a gay bar in Amsterdam, where we met. He was 30s, from Utrecht, into leather. We went back to my hotel room just after 1 a.m. and after sharing a joint, went to bed...

First we kissed, then I sucked him and he then sucked me and after that we moved into a ‘69’ position and he came in my mouth. Then I moved to fuck him and he put a condom on me with KY over it. I entered him (him sitting on me) and he was wanking himself, and we started using poppers and I then came in him and then he came over my chest. After pulling out, we kissed for a while then went to sleep.

Encoding of Account
(PARTNER LIST) C3, 30s, male, Utrecht, into leather

(encoded string of sex act sequence):

MDK AS PS MS,NM (AE,CN/1H,WN)/p MDK

The extended natural language description of the session is thus reduced to the code string (last line), used in all analysis programs.

There are several ways of now performing such a matching; one consists of defining a Levenshtein distance between two sequences (Sankoff and Kruskal 1983:18 et f), but in this instance a related method was used which was developed in molecular biology for sequence homology searches (Feng et al. 1985). These methods produce a score by which a match could be said to be made if it exceeds a stipulated limit. In brief, the sequence of sexual acts making up a session of the reference subject should be the same as the potential partner’s session if it is to match. There are some difficulties and provisos which must be made. First, the respective modalities for any act must correspond rather than be identical (e.g., if the reference subject is active for a given act, the partner must be passive for it). Moreover, allowance had to be made for “chunking”: some subjects make finer distinctions than others in describing a sequence. Finally, a “rarity” weighting was applied; some acts (e.g., “fisting,” ano-branchial insertion) are very infrequent, and their occurrence therefore has a higher surprisal value. Their occurrence is deemed more important in matching than common acts such as masturbation.

When comparison was also made of the content of the strings, the rate of matching reduced considerably; now only two percent of strings are matched as referring to the “same” session.

Further descriptive information in the diary entries (e.g., the occupation of the partner, or whether drugs were used) can tighten up yet further the likelihood of partner identification, though this has not yet been done to any considerable extent.

2.2.2.4. Evaluating “Anonymous” identification

At this juncture, little more can be said about the validity of the process: the matching certainly has face validity, but we cannot know that (for this data set) the match is correct, since it is a totally anonymous (unidentified) set of data. The next stage of validation will consist of testing whether the procedure can identify the (known) real partners from their session strings. In the meantime an interesting and potentially very important hypothesis arises concerning the “class of potential partners” in a known sexual session. Quite independently of the question of how to find the “real” partner—i.e., assuming that s/he is in this class—how similar or homogeneous is the subset of candidates for being the partner? Do those engaging in the same (or structurally identical, or similar) sequences of sexual behavior resemble each other in other ways? In particular, is their pattern of other partners similar? If so, this will provide confirmation of the “like-me” characteristics of other social networks, and incidentally provide aggregate estimates of mixing for the Anderson models.24

In the meantime there are two ways in which this technique can lead to improved and more extensive matching:

- “Coverage” should be as high as possible: even when a match occurs there is a possibility that it is fallacious in the sense that partners outside the diary sample were actually involved and the temporal and behavioral matching was therefore purely coincidental. With higher proportions of the sexually active actually involved, this probability is decreased.
- The date-limits of the diaries, should be as close to identical as possible, i.e., if most respondents are completing their diaries over the same period, there is again a higher probability of correct matching.

Current research on the reliability and validity of the diary method and its use in this context is also helping us to understand how and when two scripts or session-strings are to be considered identical.

3. CONCLUSIONS

In Project SIGMA we have tried at all times to keep the atom of social networks—the dyad—as our main unit, rather than the individual. Most obviously this means looking at (non-solitary) sexual behavior as being primarily the outcome of an interaction rather than as an individual propensity, as processes such as negotiation then become central to explaining what
Social and Sexual Networks

is occurring epidemiologically. The recent debate on so-called "relapse" provides a good instance of the difference that this makes in interpretation.

But the larger issues of networks also loom large, and we have to admit mixed success. This chapter has therefore been written in a self-critical mode, so that others can learn from our mistakes and successes. Given the unwillingness of funding agencies to underwrite some of the more ambitious proposals (very possibly with good reason), we have had to take a more pragmatic line than perhaps we would have wished. But the study has also shown us that a surprising amount can be done within the confines of a fairly conventional longitudinal study to examine network characteristics. For example, examination of triads of sexual and other interactions allows direct estimates to be made of transitivity and other "bias" parameters even if we know little about the detailed topology of the entire network. But equally there are some matters which, however difficult and expensive, can only be tackled by a direct, full-blown network methodology. It will be a tragedy if such work has to be abandoned in favor of a purely individual-based mode of enquiry and analysis which cannot represent these crucial aspects of transmission.

NOTES

1. I am grateful to the Department of Health and to the Medical Research Council funding which supports the work reported, and views expressed are my own and do not necessarily represent those of the funding authorities. I am also especially grateful to Dr. Christopher Joyce (now Research Scientist, AIDS/HIV Division, Communicable Diseases Research Centre, Colindale, London) for his collaboration and work presented in section 2.2 of this paper.

2. The acronym SIGMA represents Socio-sexual Investigations of Gay Men and AIDS. Project SIGMA is a a longitudinal, non-clinical based, ethological and behavioral study of the sexual and social lifestyle of gay and bisexual men in England and Wales. (It is also part of the English study under the auspices of WHO Global Programme on AIDS Homosexual Response Studies).

SIGMA is one of the largest cohort studies in Europe and the only study in the U.K. to have emerged from the gay community. Initial work began in 1983, and funding followed in 1987. To date, the Project has interviewed over one thousand men, half of whom have been interviewed four times at [median] intervals of ten months. The main aims of the study are to describe the sexual behavior and lifestyles of gay and bisexual men; to monitor changes in sexual behavior in relation to HIV/AIDS; to examine attitudes to different sexual behaviors and relationships; to investigate reactions to safer sex practices; to estimate prevalence of HIV and other viral infections in a non-clinical group of gay and bisexual men.

Project SIGMA uses several complementary methods of obtaining information, including:

- The detailed structured interview in which each respondent is asked for detailed information on sexual history and current practices (centered on the Index of Sexual Behavior (Coxon et al. 1992); numbers and characteristics of sexual partner, health, and attitudes towards HIV and safer sex.

- The sexual diaries (Coxon 1989) are a daily record of sexual activity kept by respondents for twelve months after each interview. So far we have collected information on thirty thousand sexual encounters which allows a unique analysis of their structure.

- Blood and/or saliva samples are also collected at the interview by trained staff and tested for HIV-1 antibodies and other viral markers. Results are available to respondents through trained counselors.

- The postal survey of sexual behaviour is a self-completion questionnaire which appears in the gay press periodically.

3. "Close-knitness" (although not Rapoport's term) is measured by the average rate of acquisition of new contacts, and the number of steps (links) necessary to reach excom- putation—the latter referred to as the "ultimate connectivity" the fraction of the population finally reached.

4. The two factors of Age and Relationship-type were dichotomized and crossed to produce the nine-fold typology: Age was split at twenty-one (the age of homosexual contact in England and Wales) and thirty-nine (after which men would have grown up when any homosexual behavior was illegal, before the 1969 Sexual Offences Act), and Relationship-type into Closed, Open and No Regular Relationship.

5. At a later stage we decided that the anonymity undertaking might be a cause of shielding ourselves in the methodological foot (see the discussion in Coxon et al. 1993), but there were (and are) excellent reasons why gay men need to be persuaded that such information is safe and cannot be used against them.

6. Defined as the old county boundaries of the County of Glamorgan, but perhaps more accurately expressed as "Cardiff and the surrounding area."

7. "Parent" (sibling) bias:

Pr [xRy | yRx], and

"Sibling" (b) bias:

Pr [xRy | (x, zRx zRy)]
These combine with the actual outdegree ("axone density", a) to form the "Reduced axone density", a), by the following expression:

a = a - p - (a - 1) α

8. A. A. "7courage" (USA) and "7ooper" (Australia), i.e., public toilets. Sub-projects of SIGMA and of the St. Mary's Paddington group studied this group in various locations—though many aspects were at variance with Humphreys' (1979) conclusions, especially about the considerable preponderance of heterosexual married men frequenting the cottages (though they are a significant fraction) Although the sex which takes place is by and large safe, there are particular (often more deserted) ones where entirely unsafe sex is the norm.

9. Terminology referring to the modality of sexual behavior is confusing (see Coxon 1989; Coxon et al. 1992). Certain medical usage is "insecure" vs. "active", "passive" vs. "recipient" (but this is useless for non-penetrative sexual behavior, such as masturbation); and "donor" vs. "recipient" inas much as it is unusable and in common parlance among gay men. Following linguistic usage, the active partner does X to the other partner whilst the passive partner lets X done to him by the other partner. Note that "inactive" is not necessary equivalent to the "active" partner.

10. Present research suggests that it may well be highly infectious at particular stages, especially after infection and before sero-conversion; more worryingly, the same may be true of fellatio at this stage (see Koopman et al. 1992).

11. I.e. the "never fuck" (0 link) is often more stable over time than the other individual roles (Coxon et al. 1993).

12. This shortcoming and ways to attempt to overcome it are described in Section Three.

13. Respondent fifteen is P as regards anal intercourse with men, but has active vaginal intercourse with a partner in London.

[234] Social and Sexual Networks

15. At later waves, this prohibition was relaxed, though not to any considerable extent, and it is still under discussion in the Project.

16. Together with information about the partnership—how long known, sexual in/exclusivity, number and sex of any other partners of the partner, and frequency of seeing.

17. The sexual diary data used were those obtained from the first Gay Times cohort, collected in 1986 and documented in Conron 1990. The total data-set refers to 188 individuals, who performed 7332 sexual acts in an average of 23.4 sessions in the month.

18. The problem is not unusual in any matching context, but five characteristics of this sort are not usually enough fully to disambiguate different candidates.

19. A "sexual session" in the structure of sexual behavior developed by Project SIGMA (Conron et al. 1992) is the main unit of such behavior; the "sentence" of sexual activity—self-sufficient and intrinsically well-formed. It occurs at a given (specified) time and place, consists of one or more sexual acts, and involves at least one, usually two (and sometimes more) people. It is typically terminated by sleep, a non-sexual intermission, or change of partner. The characteristics of a Sexual Session include four components: the Setting, the Procedure [e.g., preluditory drink, nixits], the Accompaniments ["tools," leather], and the Partner Specification mentioned above, and the main content is the sequential specification of the components' sexual acts.

20. This figure plus or minus five percent holds for all the sexual diary data sets.

21. Which basically estimates the number of insertions and deletions necessary to turn one string into another.

22. A term borrowed from information theory (Luce 1960), where the "surprisal" of an act is inversely proportional to the frequency of its occurrence: \( h_i = \log_2 p_i \).

23. A total of 344 possible matches involving 298 individuals—i.e., a possible matching of 149 pairs of individuals from a total of 166 diaries.

24. Similar concerns, with a different methodology, have been expressed in the context of research on AIDS in Africa by Orubuloye, Caldwell, and Caldwell 1992.

Appendix One

Volume of Direct and Indirect Paths Between Points in Figure 1

<table>
<thead>
<tr>
<th>Pts:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Culture and Sexual Risk
Anthropological Perspectives on AIDS

Edited by

Han ten Brummelhuis
University of Amsterdam
The Netherlands

and

Gilbert Herdt
University of Chicago
Illinois, USA

1995
Chapter

TEN

Risk in Context:
The Use of Sexual
Diary Data to
Analyse Homosexual
Risk Behaviour

Anthony P.M. Coxon and N.H. Coxon

This is a study of gay men's sexual behaviour under the impact of AIDS which uses the method of diaries. It is an integral part of the work of Project SIGMA (Socio-Sexual Investigations of Gay Men and AIDS), which is a longitudinal study of the sexual and social lifestyle of gay and bisexual men in England and Wales (and also part of the English study under the auspices of WHO Global Programme on AIDS Homosexual Response Studies). SIGMA is one of the largest cohort studies in Europe and the only study in the UK to have emerged from the gay community. Initial work began in 1983, and funding followed in 1987. To date, the Project has interviewed over 1000 men, half of whom have been interviewed four times at (median) intervals of 10 months. The main aims of the study are to describe the sexual behaviour and lifestyles of gay and bisexual men; to monitor changes in sexual behaviour in relation to HIV/AIDS; to examine attitudes to different sexual behaviours and relationships; to investigate reactions to safer sex practices; and to estimate prevalence of HIV and other viral infections in a non-clinic group of gay and bisexual men.

Project SIGMA uses several complementary methods of obtaining information, including: the detailed structured interview in which each respondent is asked for detailed information on sexual history and current practices (centre upon the Index of Sexual Behaviour, Coxon et al. 1992), numbers of sexual partners, health, and attitudes towards HIV and safer sex; the sexual diaries, a daily record of sexual activity kept by respondents for a month after each interview (see Coxon 1988b; so far we have collected information on about 30,000 sexual encounters which allows a unique analysis of their structure); blood and/or saliva samples collected at the interview by trained staff and tested for HIV-1 antibodies and other viral markers (results are available to respondents through trained counsellors); and the postal survey of sexual behaviour, a self-completion questionnaire which appears in the gay press periodically. In this chapter we concentrate primarily upon the method of sexual diaries.

METHODS FOR COLLECTING SEXUAL DATA

There are various methods for collecting data for analysis of sexual behaviour, and their merits depend on what is needed from the data. Retrospective interviews and questionnaires are useful for obtaining general and 'memorable' aspects of behaviour (such as whether a given sexual behaviour has ever been experienced), but they are not efficient means of obtaining detailed data (such as accurate estimates of the frequency, or sequences of behaviour) that can more appropriately be obtained from the method of sexual diaries. The main reason for this is that accurate retrospective recall of detailed sexual behaviour is rarely reliable or accurate beyond a few days.

On the other hand, selection bias is higher in the use of diaries than in survey techniques, simply because keeping a monthly diary on a daily basis involves a more prolonged commitment, and it is unlikely to be undertaken by those with a very inactive sex life, thus biasing estimates upward. A further possible drawback is that volunteers may record their behaviour in a way which they expect it should be (perhaps modifying the account to be more in accord with safer sex). On investigation, these drawbacks seem not to appear, and many results agree with one's intuition and with the data obtained by different methods. Preliminary research evidence suggests that
SIGMA sexual diary form

ID Number : XY / 00123 / 5
Week beginning: Sun 3 / May / 1992

Remember, each session should include:
- The Time, The Place, The Partners (from partner list)
- Then, the session in your own words (or the code if you are confident).
- If you 'come' (ejaculate) in the session, remember to be explicit about
  where it goes and always to record the use of condoms.
- List any accompaniments you use (poppers, lubricants, drugs, sex toys, ...)

---

SUN DAY
9 am My flat, P1
We deep kissed, and moved into a '69.' Whilst doing it I began to finger him. Then he wanked me (both using poppers) and I came. Following that I wanked him till he came.

3rd

MON DAY
7.30 am I woke up to find P1 wanking me. Then he sucked me off, and I came in his mouth. We began using poppers and I sucked him, carrying on to fuck him (with condom) whilst he wanked himself. He came, I didn't.

---

Figure 1. Part of a SIGMA diary.

counts of sexual behaviours based upon self-reports in the interview
are systematic distortions of the more reliable counts derived from
the diaries.3

COLLECTING DIARY DATA IN PROJECT SIGMA

At each wave of the investigation, the subject produces a (retrospective) diary of the last week's sexual behaviour in the presence of the interviewer in the context of the ordinary interview. This ensures that the diarist understands what needs to go into the diary, and also provides indirect evidence of the ability to recall detailed information. At the end, he is given a month diary kit (SIGMA 1993), and he returns the completed diary after the month period. The diary form is completed each day, and explicit instructions are given for its completion. The diary is filled out in ordinary language, within the framework specified, and respondents are encouraged to use 'street language' if they wish to. An example of part of a completed week form is given in Figure 1.

When the completed diary is received back, it is encoded according to a scheme described in Coxon et al. 1992, summarised, and then entered into a database for subsequent analysis.

The chief methodological advantages of the diary method are its ease and flexibility: (a) it is naturalistic (some respondents are accustomed to keeping some sort of sexual diary anyway); (b) it is easily adapted to record sexual data from those of any sexual orientation (or, indeed, to domains other than sex); (c) it can be augmented to obtain other concurrent information (such as alcohol and other substance consumption in a sexual context; see Weatherburn et al. 1993); and (d) it can be expanded to take into account such aspects as coital position.

There are a number of more substantive advantages of the diary method: (a) it allows the detail of sexual behaviour to be recorded and analysed, and to be encoded intact, since the researcher knows the context of any sexual act — location (time, place), the actors (sexual partners), the adjuncts to sex (such as condoms, lubricants, toys, 'poppers'), and is able to place their use in the correct context; (b) uniquely, it permits the analysis not only of individual behaviour (such as, 'What is the average number of times a gay man engages in oral intercourse in a month?' but of the 'volume' (or 'output') to use Kinsey's phrase) of behaviour, where the denominator is all the sexual acts involved (as in, 'What proportion of young gay men's sexual activity in a month consists of oral intercourse?'); (c) it allows sequences of sexual activity to be analysed. Sometimes this can be crucial, for we need to know whether unsafe behaviour tends to occur at the start, middle, or end of a sexual session, and the meaning of acts such as anal intercourse can be substantially different when it occurs as an end-marker rather than as an incidental activity; and (d) it allows issues of power or dominance in sexual behaviour to be investigated by systematically encoding the morality of the act (which partner does what to whom) and thus allowing gendered distinctions to be made between reciprocated and dominance sessions — the former where the other partner tends to do the same thing in return to his partner, the latter where one partner is repeatedly the submissive, the other the dominant partner (Davies 1990).

In this chapter, these claims are illustrated with respect to the analysis of two aspects of sexual behaviour, and especially anal intercourse: the effect of relationship type on sexual behaviour and behavioural progression in a sexual act (answering the questions: where is anal intercourse located in a sexual session, and what difference does it make?)
BACKGROUND

Before proceeding further, it is important that the reader understand the basics of the system of encoding the structure of sexual behaviour that we have developed, since the results in the tables make use of it (Coxon 1988a; Coxon et al. 1992). Essentially, every diary is encoded into a database from which application programs may analyse the results. The encoding process simply translates the respondent’s description of sexual behaviour into a structured formal language that readily lends itself to computer representation.

Every diary pack has a preliminary face sheet eliciting information on HIV status, relationship type, geographic location, etc., in the form of a few preliminary questions, together with a table for the respondent to list sexual partners (with their characteristics) who are involved in their sexual transactions. The characteristics include demographic and descriptive information about each partner, and the list is built up as the diary progresses.

The remainder of the diary looks similar to a normal diary, giving space for every day in which the diarists may write what they have done. They are instructed to enter the time and date, then to describe in normal (‘street’) language precisely what they did, according to a provided set of criteria. Each such spatially and time-limited sexual encounter is referred to as a session and may involve one or more persons. Each session forms a separate record in the database. The sexual session is subdivided into sexual acts, each of which is finally described by the behaviour and its modality and by its outcome — basically whether and how and where ejaculation occurs.

To aid understanding, take the following simple example of a sexual session:

\[ \text{Session: } = [ \text{PW AS AF} ] \]

This session consists of three acts (a session may contain many acts, but must contain at least one). The first character of each act is the modality, and any remaining characters are the behaviour. (The encoding process is far more complex than this, but it is not necessary to go into depth for the results at this stage.)

The behaviour describes what (sexual act) is done, and the modality describes who does it. There are five modalities, which define who does what and to whom from the diarist’s (also referred to as Ego’s) point of view. Ego’s partner is referred to as Alter (see Table 1).

A comment is in order about the ‘I’ modality. Although Ego is not involved directly in this modality, what his partner does by himself may have possible consequences for HIV transmission (the original stimulus for our development of the method), such as when Alter masturbates himself and ejaculates over Ego.

In the example above, ‘PW’ means that the behaviour here is ‘W’ (masturbation), and the modality is Passive, i.e., Ego is masturbated by Alter. The whole sexual session quoted would be read as:

Ego is passively masturbated (PW)
then Ego actively performs fellatio on Alter (AS)
then Ego actively engages in anal intercourse with Alter (AF)

In the actual diary this may well have been written as:

He wanked me, then I sucked him off, then I fucked him.

The diary thus preserves the sequence or progression of behaviour, which is an important aspect of many types of analysis. Table 2 lists the current set of recognized behaviours (sexual acts).

<table>
<thead>
<tr>
<th>Modality</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>Behaviour performed by Ego on himself</td>
</tr>
<tr>
<td>Active</td>
<td>Behaviour performed by Ego on his partner</td>
</tr>
<tr>
<td>Passive</td>
<td>Behaviour performed by his partner on himself</td>
</tr>
<tr>
<td>Mutual</td>
<td>Behaviour performed simultaneously by both partners to each other</td>
</tr>
<tr>
<td>Him</td>
<td>(Opposite of ‘Self’). Ego’s partner does it to himself</td>
</tr>
</tbody>
</table>

A GENERAL VIEW OF BEHAVIOUR

At the simplest level, information can be obtained from respondents’ diaries to match commonly used interview-based variables such as information referring to the prevalence and incidence of various sexual activities. In interviews, this is obtained by asking the respondent him/herself to estimate how often (or whether) a given sexual act has been engaged in during a specified period. In the sexual diaries such information is derived post factum by the researcher by counting how often a given activity occurs in the diary script.
The first area of interest is to see how figures derived from sexual diaries compare at the aggregate level to interview responses. Table 3 deals with incidence of sexual behaviour by modality, i.e., the percentage of those (in a month) who had ever in the course of their diaries done each of several given sexual behaviours (rows) in the given modality (columns). Thus, 24 per cent of the 610 diarists had engaged in active anal intercourse ('fucked a guy') and 25 per cent had engaged in passive anal intercourse ('been fucked by a guy') in the last month.

The percentages for Self and Him modalities are generally small or not relevant for our present purposes, due to their physical difficulty or impossibility or irrelevance. There are some exceptions: Solo Masturbation is important because it is the most prevalent sexual behaviour, and His Masturbation indicates that a quarter of the diarists experienced their partner's solo masturbation in a sexual session.

The more interesting point concerns active and passive variants of sexual behaviours. In a stable and closed population, the figures for active and passive variants of a sexual behaviour should be identical. In fact, the figures are surprisingly close to equality, and there is no tendency for there to be more active diarists than passive diarists, as one might expect. For sexual diary data the differences are smaller than the data from any other method, indicating their greater validity.

How similar are these diary data to aggregated self-report data obtained by the interview method? At first sight and looking at the actual percentages, they are rather different, with some major discrepancies: the rank-order correlation is modest ($r = 0.77$). More interestingly, the Pearsonian (linear) correlation is much higher ($r = 0.95$), and the regression equation

$$Y = 15.82 + 0.94X$$

reveals that the relationship is close to absolute, but with an added constant of about 16. That is, if the diary data are more reliable, then the interview method systematically overestimates the incidence figures but keeps a virtually identical profile. One possibility is that the interview respondents actually use an interval greater than one month and correctly remember having done $X$, but have done it in a longer period. There is one notable exception: the figures for anal
intercourse (active: 29 vs. 24 and passive: 28 vs. 25) are markedly similar, which may indicate that gay men are able to recall with unusual accuracy whether or not they have engaged in this most implicated behaviour.

**ANALYSIS OF 'VOLUME': BEHAVIOURAL TRENDS**

In a report on an earlier diary sample (Coxon and Carballo 1989) it has been said that 'the vast majority of gay men's sexual activity is taken up by the three main behaviours of masturbation, fellatio and anal intercourse.' The justification for this statement is clearly seen in diary analysis, where the act (as opposed to the person) can be used as the unit of counting. This is termed 'analysis of volume' in our account and is a mode of analysis only feasible using the diary method. It makes it possible to ask not only how many men did X, or how often they did it, but also: 'What proportion (as a fraction of all sexual acts performed) of a person's (or a group's) total outlet consisted of X?' Table 4 illustrates the proportions (of the total number of sexual acts in this data set) which various sexual behaviours occupy.

In fact, over 80 per cent of all behaviour in the sample consists of these three main activities: masturbation, fellatio, and anal intercourse. Masturbation is by far the most common activity and accounts for three-fifths of the total outlet, with Solo Masturbation accounting for over one-third of all sexual acts. The low figure for DK (deep kissing) is simply a reflection of the fact that most diarists don't consider it to be a sexual act, and therefore don't record it consistently. The next 16 per cent of sexual outlet is distributed between BR (frottage), DK (deep kissing), FG (digital-anal insertion), MA (massege) and RI (anilingus).

Expressed in terms of the total sexual outlet for the group, heterosexual activity (vaginal intercourse (VF) and cunnilingus (CN)) is very low indeed. Table 3 shows that 3 per cent of the individuals in the last month had a heterosexual encounter, but this represents merely 0.4 per cent of the total sexual outlet. It would seem that Britain is peculiar in this respect compared to other WHO Homosexual Response Studies sites (see Coxon 1992: §3.2), showing the lowest figures out of all the seven nations.

**BEHAVIOUR AND RELATIONSHIPS**

Within the diaries, sexual relationship types are defined, following the SIGMA conventions, as: (a) one exclusive regular partner

<table>
<thead>
<tr>
<th>Behaviour</th>
<th>Frequency of acts</th>
<th>Percentage of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masturbation (W)</td>
<td>20,976</td>
<td>58.3</td>
</tr>
<tr>
<td>Fellatio (S)</td>
<td>5,990</td>
<td>16.7</td>
</tr>
<tr>
<td>Anal int. (F)</td>
<td>2,017</td>
<td>5.6</td>
</tr>
<tr>
<td>Frottage (BR)</td>
<td>1,683</td>
<td>4.4</td>
</tr>
<tr>
<td>Deep kissing</td>
<td>1,346</td>
<td>3.7</td>
</tr>
<tr>
<td>Ano-digital insertion (FG)</td>
<td>1,057</td>
<td>2.9</td>
</tr>
<tr>
<td>Massage</td>
<td>921</td>
<td>2.6</td>
</tr>
<tr>
<td>Anilingus (RI)</td>
<td>813</td>
<td>2.3</td>
</tr>
<tr>
<td>Inter-femoral frottage (TF)</td>
<td>460</td>
<td>1.3</td>
</tr>
<tr>
<td>Corporal punishment</td>
<td>269</td>
<td>0.8</td>
</tr>
<tr>
<td>Nipple play (TT)</td>
<td>148</td>
<td>0.4</td>
</tr>
<tr>
<td>Ano-brachial insertion (Fi)</td>
<td>133</td>
<td>0.4</td>
</tr>
<tr>
<td>Vaginal intercourse (VF)</td>
<td>126</td>
<td>0.4</td>
</tr>
<tr>
<td>Lindinism (WS)</td>
<td>39</td>
<td>$^{0.1}$</td>
</tr>
<tr>
<td>Cunnilingus</td>
<td>7</td>
<td>0.02</td>
</tr>
<tr>
<td>Total acts</td>
<td>35,981</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Table 5. SIGMA Typology: Relationship by Age Types**

<table>
<thead>
<tr>
<th>Age</th>
<th>Under 21</th>
<th>21 to 39</th>
<th>Above 39</th>
</tr>
</thead>
<tbody>
<tr>
<td>One regular partner</td>
<td>I</td>
<td>II</td>
<td>III</td>
</tr>
<tr>
<td>Regular partner(s) and others</td>
<td>IV</td>
<td>V</td>
<td>VI</td>
</tr>
<tr>
<td>No regular partner</td>
<td>VII</td>
<td>VIII</td>
<td>IX</td>
</tr>
</tbody>
</table>

('closed'); (b) one or more regular partners and others ('open'); and (c) no regular partner. Each of these relationship types is subdivided into three age ranges. Each cell of the resulting nine-fold design is given a label, in the form of a roman numeral, as shown in Table 5.

Patterns of sexual behaviour can then be studied by looking at how the sexual outlet of each of these nine types differs. This is done
by treating that type's total sexual outlet as 100 per cent, and then looking at how it is divided between the three most common sexual behaviours and their associated modalities. (This is termed unconditional or volume analysis, as opposed to conditional analysis, where the individual is the unit of analysis.)

An exhaustive analysis of the whole data set by type was performed and gave the results presented in Table 6. They are worthy of considerable attention, since once again, these are data which are unobtainable by any other method. Each section of the table represents a separate modality, which is then in turn divided by sexual activity. For example, Table 6.1 indicates that 22.2 per cent of the total 694 (Table 6.5) sexual acts done by type I individuals (those aged 21 and under and in a closed relationship) consist of solitary masturbation, compared to 54.1 per cent of type VII (those aged 21 and under with no regular relationship). This suggests that having no regular partner has the effect of doubling the reliance on masturbation for these young men (see Davies et al. 1992).

The behavioural characteristics for each type are well defined:

**Masturbation.** Those with no regular partner (VII, VIII, IX) show a very high incidence of solo-masturbation, whereas those in closed relationships (I, II, III) have a very low incidence.

**Anal intercourse.** Those with no regular partner (VII, VIII, IX) show the lowest incidence of any form of anal intercourse, with most of their values being significantly less than the 'average' figure indicated in the ALL column. Those in closed relationships show a consistently higher level of both active and passive anal intercourse. Active anal intercourse is significantly higher for type III (the older partners in a regular relationship), and this is shown in Table 7. But there is no significant variation in the passive form (see Coxon et al. 1993 for further treatment of this question).

**Fellatio.** Fellatio seems to transcend the relationship limits, and does not show any consistent relationship pattern. Types IV, V, and VI have the highest figures for fellatio, but we should bear in mind that type IV is by far the smallest set, and so we cannot treat the result as reliable.

No other form of behaviour exhibits a significant relationship type or age dependency.

**BEHAVIOURAL PROGRESSION WITHIN THE SESSION**

Anal intercourse is recognized as the form of sexual behaviour that is implicated most significantly in the transmission of HIV. From a

---

### Table 6. 'Volume' Analysis of Sexual Behaviour by Modality and Type of Relationship (the following percentages are expressed as a proportion of the total sexual outlet of the indicated type of sexual behaviour:

- W = masturbation; F = anal intercourse; S = fellatio)

#### 6.1: Modality = self: solo activity by act (masturbation) and type

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>VI</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>34%</td>
<td>22.5</td>
<td>22.8</td>
<td>19.7</td>
<td>15.3</td>
<td>31.0</td>
<td>21.4</td>
<td>54.1</td>
<td>53.7</td>
</tr>
</tbody>
</table>

#### 6.2: Modality = active: active (Ego to Alter) by acts and type

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>6.9</td>
<td>10.0</td>
<td>7.5</td>
<td>7.7</td>
<td>2.4</td>
<td>7.2</td>
<td>9.9</td>
<td>4.2</td>
<td>4.8</td>
<td>6.3</td>
</tr>
<tr>
<td>F</td>
<td>2.6</td>
<td>4.0</td>
<td>2.8</td>
<td>10.5</td>
<td>2.4</td>
<td>2.4</td>
<td>3.5</td>
<td>1.4</td>
<td>1.2</td>
<td>2.6</td>
</tr>
<tr>
<td>S</td>
<td>7.5</td>
<td>7.2</td>
<td>7.0</td>
<td>7.5</td>
<td>12.9</td>
<td>8.0</td>
<td>10.3</td>
<td>6.4</td>
<td>5.8</td>
<td>6.6</td>
</tr>
</tbody>
</table>

#### 6.3: Modality = passive: (Alter to Ego) by acts and type

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>6.7</td>
<td>8.5</td>
<td>7.5</td>
<td>5.2</td>
<td>7.0</td>
<td>7.1</td>
<td>8.4</td>
<td>5.0</td>
<td>4.1</td>
<td>5.0</td>
</tr>
<tr>
<td>F</td>
<td>2.4</td>
<td>3.0</td>
<td>4.0</td>
<td>3.8</td>
<td>4.7</td>
<td>2.2</td>
<td>2.2</td>
<td>1.5</td>
<td>1.5</td>
<td>2.2</td>
</tr>
<tr>
<td>S</td>
<td>6.2</td>
<td>7.5</td>
<td>7.0</td>
<td>3.6</td>
<td>9.4</td>
<td>7.3</td>
<td>5.2</td>
<td>4.8</td>
<td>5.0</td>
<td>4.2</td>
</tr>
</tbody>
</table>

#### 6.4: Modality = mutual: (Alter and Ego simultaneously) by act and type

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>7.4</td>
<td>7.2</td>
<td>6.0</td>
<td>3.6</td>
<td>11.8</td>
<td>6.8</td>
<td>10.4</td>
<td>6.8</td>
<td>5.6</td>
<td>11.0</td>
</tr>
<tr>
<td>F</td>
<td>2.8</td>
<td>2.6</td>
<td>3.2</td>
<td>4.8</td>
<td>4.7</td>
<td>2.4</td>
<td>4.5</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
</tr>
</tbody>
</table>

#### 6.5: Total sexual outlet by type (frequency of acts)

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>194155</td>
<td>694</td>
<td>2748</td>
<td>478</td>
<td>85</td>
<td>1136</td>
<td>2792</td>
<td>1368</td>
<td>2545</td>
<td>5800</td>
</tr>
</tbody>
</table>
behavioural point of view, it would be interesting to see what types of behaviour are most likely to precede and lead on to anal intercourse of some form. We already know that anal intercourse is most typically the end-marker of a session, and the sequence in a sexual session may be so constrained that anal intercourse becomes virtually inevitable. Moreover, different behaviours may precede the active as opposed to the passive variant. Is this true?

In a preliminary analysis of the co-occurrence of sexual acts it was found that the two acts most likely to occur in the same sexual session as anal intercourse are masturbation and fellatio. Starting with this, precedence analysis was then applied to the sessions to see what behaviour is most likely to lead on to anal intercourse.

'Precedence analysis' is essentially a frequency count of how many times a given act precedes another in a sexual session. The analysis uses a 'sensitivity' parameter called 'width' which defines how many acts are taken into consideration by the analysis as preceding the reference act. For example, consider the following session:

**Session 1:** { AW PW PF AS }

Let us use PF as the reference base; the AS is ignored in a precedence analysis, since it follows the reference. If the width is one, then only PW is counted as the precedent, but if the width is 2, then both PW and AW count as precedents. Thus the width value defines how many acts are visible to the analysis previous to the reference base. But the same act can occur more than once in a session, and this has to be allowed for. Consider the following example:

**Session 2:** { AFI AS PW AS AF }

If AF is the reference base, then the three (distinct) precedent acts are: AFI, AS, and PW. But AS occurs twice, and is therefore given a higher precedence value. In the above example, the precedence value of AS (with AF as the reference) has a value of 2 if the width is 3, but a value of 1 if the width is 2, the reason being that the second AS is out of the range of width visibility when the width is 2. (Width does not include the reference act itself.)

Let us turn now to the diary data (see Table 8). Looking at both the Active and Passive tables, the rank order of likely precedence is the same, independent of the width of the precedence. Active/Insertive anal intercourse is most likely to be preceded (in order) by:

(Passive Sucking/Receptive Fellatio)
(Active Sucking/Insertive Fellatio)
(Active Wanking/Masturbation), and
(Passive Wanking/Masturbation)

In brief, this can be expressed as a rule: 'If I fuck a guy, sucking is most likely to precede it (rather than wanking), and I'm most likely to be sucked

---

**Table 7. Sexual Outlet for Active Anal Intercourse (AI)**

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Under 21</th>
<th>21 to 39</th>
<th>Over 39</th>
</tr>
</thead>
<tbody>
<tr>
<td>One partner</td>
<td>4.0</td>
<td>2.8</td>
<td>10.6</td>
</tr>
<tr>
<td>Regular partners and others</td>
<td>2.4</td>
<td>2.4</td>
<td>3.6</td>
</tr>
<tr>
<td>No regular partner</td>
<td>1.4</td>
<td>1.2</td>
<td>2.7</td>
</tr>
</tbody>
</table>

**Table 8. Precedence Analysis for Anal Intercourse**

<table>
<thead>
<tr>
<th>Act</th>
<th>Width</th>
<th>Width</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>AS</td>
<td>49</td>
<td>89</td>
<td>126</td>
</tr>
<tr>
<td>PS</td>
<td>85</td>
<td>118</td>
<td>146</td>
</tr>
<tr>
<td>AW</td>
<td>31</td>
<td>61</td>
<td>73</td>
</tr>
<tr>
<td>PW</td>
<td>31</td>
<td>51</td>
<td>64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Act</th>
<th>Width</th>
<th>Width</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>AS</td>
<td>50</td>
<td>89</td>
<td>123</td>
</tr>
<tr>
<td>PS</td>
<td>42</td>
<td>80</td>
<td>101</td>
</tr>
<tr>
<td>AW</td>
<td>26</td>
<td>38</td>
<td>52</td>
</tr>
<tr>
<td>PW</td>
<td>33</td>
<td>51</td>
<td>61</td>
</tr>
</tbody>
</table>
by him first.’

Passive/receptive anal intercourse, by contrast, is most likely to be preceded (in order) by:

(Active Sucking/Insertive Fellatio)
(Passive Sucking/Receptive Fellatio)
(Passive Wanking/Masturbation), and
(Active Wanking/Masturbation)

In brief: ‘If I am fucked by a guy, sucking is, again, most likely to precede it, and I’m most likely to suck him first.’ (This accords with gay received wisdom and underwrites once again the symbolic dominance/subservience role of anal intercourse.)

In many ways these are the two sides of the same coin — seen from Ego’s and Alter’s perspective. A ‘successor analysis’ of active and passive fellatio would reveal the similar point: ‘If you suck your partner, you’re likely to go on to be sucked by him,’ and ‘If you are sucked by your partner, you are likely to be expected to go on to fuck him.’

If anal intercourse occurs in a sexual session, then it is most commonly the end-marker, and is very likely indeed to result in ejaculation, so this is where attention needs to be focused for a more accurate knowledge of how risk occurs (or is prevented). But this is another story (see Coxon and Coxon 1993).

USEFULNESS OF DIARY ANALYSIS

The results given by the diary analysis are very reliable and conclusive when studying the behaviour of individuals at the session level. Because of the short duration of the diaries (usually about a month or so), it is not so effective for time-generalized analysis (such as genuine ‘ever’ figures). It does, however, give a good idea of relative proportions of sexual behaviour, since total sexual outlet sets may be analysed by various independent variables (such as type, status, geographic location).

The greatest strengths of diary analysis, and areas in which it is worth particular further study, are in the analysis of the volume of sexual behaviour, of sexual role segregation, and co-occurrence, precedence, and successor analyses. These are all methods used to study behaviour at the session or individual level. The use of condoms during sessions is also an important part of the study, and one area in which diary analysis can, in principle, perform well.

Perhaps the most striking difference is that data are derived in the case of diaries, whereas data represent subjects’ accounts (or estimates) in the case of interviews. There is no warrant for believing that the two will necessarily give rise to the same conclusions (indeed, the comparison of the two is an important component of our validation studies). But it is clearly seen from this analysis that sexual diaries are capable of providing unique data which is also uniquely adapted to answering problems of considerable epidemiological and policy relevance.

USE OF DIARIES IN ANTHROPOLOGICAL SETTINGS

In the research reported here, sexual diaries have been used as an auxiliary method of data collection and validation within a Western European context. The reliability and utility of diaries in other contexts depends on a number of factors:

Whether diary-keeping is an accepted procedure in the culture. For gay men, and for sexual and romantic activity in general, diary-keeping is a popular, though often a private, pastime, and is often written in code to ensure confidentiality (McCormick 1980). The social science use (see also Pomeroy, Flax, and Wheeler 1982) consists primarily in persuading respondents to do so for scientific purposes.

Anonymity must be guaranteeable. Whilst some diarists care not who knows their secrets, for others the diary records what no one else knows, and somehow the social scientist has to safeguard this secrecy, especially if the matters being recorded are illegal, deviant, or even simply private. At least one SIGMA diarist was expelled from his lodgings because his landlady ‘found’ his sexual diary.

The diary does not depend on a written format. Although diaries are most conveniently kept in writing in literate cultures, there is no reason why they must be. We have used both tape recorders and electronic pocket memos to record (and have diarists record) their diaries.

The diary does not depend on a long period. Although we have found that a month is the minimum period for reliable data, some diarists can be persuaded to keep it for much longer. (However, we believe that persons who keep the diary longer may differ in important ways from our other respondents.) Also, card versions of day entries can be devised which allow an encrypted version of events to be sent through the mail on a daily basis, thus avoiding problems of partial-completion and loss; it also helps avoid contamination due to recognition of trends by the diarist, if this is felt to be a problem.
The diary can be completed at leisure. Communicating about sexual behaviour (and especially uncommon behaviour) is often not feasible or is subject to systematic bias if elicited in the presence of others. Diaries can be completed at any time convenient to the respondent.

Diary information can as well be based on observation as on the subject’s report. Although not always feasible, the structure of sexual behaviour underlying the diary code makes it possible to use it as a rapid shorthand for situations such as immediate re-interview, and our current studies of competent encoders’ encodings of a common video stimulus (gay male pornographic sequences) show that high agreement on ‘chunking’ and description of sexual activity can be achieved.

... ... ... ...

The sexual diary is a robust, easy-to-use method for collecting reports (and/or observations) on detailed and sequential aspects of sexual behaviour. So long as common meanings are established before its use, respondents find them straightforward and even interesting to complete, and the data can be used to answer questions which interviews based on retrospective recall are ill-adapted to answer. Probably the most telling shortcoming is that volunteer (selection) bias is likely to be strong — including the earlier point about the difference in diary-keepers. This can be minimized by careful strategies to improve participation (Rosenthal and Rosnow 1975).

Because of the inherent detail of the diary account, it lends itself to serendipitous discovery and since it provides a general trace of behaviour it can provide well-adapted information if new foci of research become imperative. With the rapid increase in our understanding of HIV transmission, this cannot but be for the best.

NOTES

1. This research is supported by grants from the Department of Health (UK), the Medical Research Council, and the World Health Organization to Project SIGMA (Essex). These bodies are in no way responsible for the views expressed in this chapter.

2. This, and other issues of reliability and validity of SIGMA sexual diary data are currently under investigation by the senior author, with funding from the Department of Health. For discussion of relevant experimental and theoretical findings see Linton (1986) and Brewer (1988).

3. Interview estimates tend to be ‘chunked’ aggregates of week estimates (e.g., a respondent will mentally estimate the last week’s frequency and then multiply it by four for a month estimate), be reported in rounded quantities (e.g., multiples of 5 or 10), and be subject to an individual distortion factor. Current research centres on the comparison of estimates provided by the subject after the return of the diary with those calculated from the diary of the same period (Coxon 1988b).

4. ‘Poppers’ refer to anyl or butyl nitrite sniffed to give a quick ‘high’ by cardiovascular dilation.

5. The respondent is asked to provide the following information about each sexual partner: [1] Whether he is regular/occasional/one-off; [2] His age (if you know it, or your guess); [3] How long you have been having sex with him; [4] Where you met him; [5] His HIV status (one of: Negative, Positive, or Don’t Know), together with the partner’s sex, if it is not male.

6. For reasons of confidentiality we did not ask the respondent to give the actual name of his contacts and in the case contacts, this was never known. Systematic network tracing is thus precluded by our adoption of this self-denying ordinance.

7. Instructions are contained in ‘Instructions for Completing a Sexual Diary,’ SIGMA 1990.

8. Since the SIGMA system was developed explicitly with HIV transmission in mind, it is necessary to distinguish which partner ejaculates, and where the ejaculate goes — in/on the partner, in a condom, etc. This is encoded in the SIGMA coding system but is not discussed further here (see Coxon et al. 1992).

9. Such a list is of its very nature open; sexual adventurousness and inventiveness, together with shifting labels for acts, make it so. Taking the respondents’ terminology as definitive does introduce methodological problems of overlapping categories, such as occur for essentially the same activities as ‘NN’ (nipple nibbling) and ‘TT’ (tit torture), and of apparently equivalent categories which reverse the modality — compare ‘sucking’ and ‘mouth-fucking’.

10. The diaries of this study are all completed by men (hence the use of male gender throughout), but there is no reason why this should be so.

11. The sexual diaries are held in DbaseIV and CARDBOX-PLUS format. The programs for subsequent analysis are contained in the software package SDA [Sexual Diary Analysis] written in TOPSPEED C by Huw Coxon with funding from the Department of Health, whose help is gratefully acknowledged. Copies of the programs, selected data, and documentation are available at cost from Project SIGMA ( attn: N.H. Coxon), University of Essex.

12. Because in general receptive anal intercourse has lower prestige or is deemed more submissive than the insertive variant.
13. Reported in SIGMA 1990 (Table 5.5, pp. 129–130). Because these data are obtained using the ISB, which has the same structure as the Diary codes, direct comparison is possible. The 17 cells entering the comparison are marked with ‘#’ in Table 3. The values, in order, are: 24 (AF), 28, 61, 58, 39, 90 (SW), 65, 63, 61, 23, 23, 10, 26, 23, 42, 38, 40 (MBR).

14. And also because it was only recently re-introduced into the list of behaviours. It was initially excluded on the grounds that it could not be implicated in HIV transmission, and was re-introduced as a marker for tracking CMV (cytomegalovirus) sero-conversion.

15. Each row of the table is better represented as an Age x Relationship table, whose entries are the relevant percentages. In previous analyses (Coxon 1987), each table is then analysed using an ANOVA or resistant analysis such as Median Polish (see note 16).

16. To illustrate the additive analysis mentioned in note 15, when the data of Table 6.2 (F) (i.e., active anal intercourse; also shown in Table 7) are analysed using Tukey’s Median Polish, the results table is as follows:

<table>
<thead>
<tr>
<th>REL</th>
<th>AGE</th>
<th>Effect of AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;21</td>
<td>21–39</td>
</tr>
<tr>
<td>Closed</td>
<td>0</td>
<td>-1.0</td>
</tr>
<tr>
<td>Open</td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>No reg.</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Effect of age:</td>
<td>0</td>
<td>-0.2</td>
</tr>
</tbody>
</table>

The overall (total) value is 2.4 per cent, and compared to this the effects of age and relationship are small — being in a closed relationship raises the percentage by 1.6 per cent; being over 39 raises it by a further 1.3 per cent, whilst having no regular relationship decreases it by 1 per cent. The entries in the body of the table are the ‘interaction’ or residual/joint effects, so the combined effect of being in a closed relationship and between 21 and 39 decreases it by 1 per cent. By far the greatest effect of all is being in a closed relationship and over 39: this adds an effect more than twice the size of the overall effect. (Adding the total, row, column, and interaction effect (necessarily) yields the original data. Thus the ‘closed relationship/age less than 21’ cell comprises:

REFERENCES


Chapter 2

The Method: Men, Diaries and Other Techniques
Chapter 2

The Method: Men, Diaries and Other Techniques

Miss Prism: You must put away your diary, Cecily. I really don’t see why you should keep a diary at all.

Cecily: I keep a diary in order to enter the wonderful secrets of my life. If I didn’t write them down, I should probably forget all about them.

*Oscar Wilde* (1895) *The Importance of Being Earnest, Act II*

Otis Dudley Duncan — himself no mean methodologist — warns his readers:

A man writing on methodology is in the same position as one who plays slide trombone; unless he is good at it, the results are more likely to interest him than his audience. (Duncan 1966, pp. 95–6)

That is a strong but justified warning, not least because this is written by someone with an emeritus chair in sociological methods! It may therefore occasion some surprise that the methodological detail of the book has been minimized in the interests of wider readability. But, however ‘spinach-like’ (to quote another of Duncan’s analogies), a methodological account is necessary if the techniques by which the information was obtained are to be reasonably assessed. In part this is because the use of systematic diary procedures are a fairly new development in studying sexual behaviour, and also because communication flow between American (US) and British behavioural scientists working in the area of behavioural AIDS research is so one-sided, as witnessed by comments from eminent American AIDS methodologists:¹

Diary approaches have not been used in any large scale AIDS-related behaviour surveys that we are aware of...

Though diaries may provide more precision than other methods, to date, diaries have not been reported to have been used in AIDS behavioral studies... (Catania 1993, pp. 151, 154)

So this chapter describes first the sampling design used to obtain subjects and then goes on to describe the method of sexual diaries, and finally describes other methods used to obtain the data analysed in this book. The methodology of the main SIGMA panel study is reported in Davies *et al.* (1993, pp. 61–81).

**Sampling subjects**

The logic behind the project sampling strategy (from which the SIGMA data-set members are drawn) is complex, but hangs on three considerations:

- issues of intensive definition (what is the population which we are trying to generalize?, e.g., are we attempting to sample ‘male homosexuals’? ‘men who have sex with men’?),
- issues of extensive definition (is it possible, at least in principle, to list the population we are referring to?), and
- issues of practicability (even if technically feasible, is it financially and practically feasible to obtain a sample?).

The answers we have argued for are as follows (Davies 1986; Davies *et al.* 1993 (pp. 66–71), Coxon 1995b).

**Intensive Definition**

*Since the study is related to HIV transmission by sexual (and possibly other) means among males, there is no need to invoke notions of self-definition as ‘homosexual’, or anything else.*
Put crudely, infection is liable to occur independently of what sexual orientation a man thinks he has and it is important to combat the notion, commonly held, of thinking of ‘homosexuality’ as an immutable and recognizable attribute not to be encouraged. Kinsey (1948, pp. 650–7) illustrates well how prevalence estimates of ‘homosexual men’ can be made to range from 4 per cent to almost 50 per cent by successively relaxing the criteria of the type of sexual contact and the time-period of sexual involvement with those of the same sex (Coxon 1988b). This is not to say, of course, that self-definitions of sexual orientation do not influence risk behaviour – in fact they do. But it is ‘sexually active men who have sex with men’ who form the notional population.

**Extensive Definition**

No extensive definition of such a population is possible (since one would have to know about the man’s sexual practices before deciding his eligibility, and issues of lying and ‘masking’ would introduce enormous biasing factors).

One has to be careful about this point; it is theoretically possible to establish a population sampling frame and, having enquired about sexual practices, then use this as a basis for selecting a sub-sample of men who have sex with men (MSM), and indeed recent national surveys (Johnson, Wadsworth et al. 1994 in the UK; Laumann, Gagnon et al. 1994 in the USA) have attempted precisely this. Considerable doubts remain about identification and estimates of the prevalence of ‘homosexual men’ by this means since there are good reasons to suppose that estimates are severely biased downward, though equally Kinsey’s estimates are almost certainly biased upwards. In any event, no general population survey of sexual behaviour was then envisaged in the UK and the cost of attempting to sample randomly on a two-stage basis (initially ‘combing’ to produce a population frame, and secondly sampling within it) was well outside our funding agencies’ costings. Moreover, after Mrs doen’t want intervention to prevent governmental funding of the main stage of such an investigation and consequent funding from private sources (the Wellcome Foundation), this has meant that the data are not accessible to other researchers.

**The SIGMA typology**

For these reasons, a different approach was attempted: to construct a design which would use the factors known to affect most the variations in homosexual men’s sexual behaviour (age and relationship-type) and to use a sampling technique (chain-link or ‘snowball sampling’) to access respondents. In their classic study Bell and Weinberg (1978, pp. 129–39) had shown that gay men’s relationships could be typified as: close-coupled, open-coupled, functional, dysfunctional and asexual and that this had a major influence on a whole spectrum of sexual behaviour and lifestyle issues. (The last three types had more to do with sexual and other problems in relationships, and were supplanted by ‘No regular relationship’.) At a later stage this categorization was made finer and became:

1. Closed/Monogamous/Exclusive: The subject is assumed to have only one exclusive partner, whether or not they live together.

2a. More than 1 regular partner: The subject is assumed to have more than one regular sexual partner but no casual partner/s.

2b. 1 regular partner and casual/s: The subject is assumed to have only one regular sexual partner, and other casual partner/s.

2c. More than 1 regular partner and casual/s: The subject is assumed to have more than one regular sexual partner, and other casual partner/s.

3. No regular sexual partner: The subject has no regular partner although he has casual partners.

In England and Wales (Scotland has a different legal system) the age variable also acted as a surrogate for English legal history. The youngest age group consists of those whose sexual activity was illegal (in
England and Wales), being under the (then) age of homosexual consent of twenty-one, and the oldest group consists of those who have grown to sexual maturity in the period before the Sexual Offences Act of 1967 when all homosexual activity (and hence, a fortiori, male sexual relationships) were illegal.

To define the typology, the three categories of sexual relationship type, closed ('monogamous'), open (at least one regular partner and others) or no regular partner, are crossed with the variable of age (categorized as under 21, 21 to 39, and over 39). These nine SIGMA relationship types are then labelled using the roman numerals I through to IX:

<table>
<thead>
<tr>
<th>Relationship Type</th>
<th>Age:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closed ('monogamous')</td>
<td>Under 21</td>
</tr>
<tr>
<td>Open ('one regular and other partners')</td>
<td>21–39</td>
</tr>
<tr>
<td>No Regular Partner</td>
<td>VI</td>
</tr>
</tbody>
</table>

It is important to stress that the respondent's identification of his relationship-type was taken as given. In some cases this would not be the same as his partner's (in particular, a relationship defined as 'closed' or exclusive by one partner might be defined as 'open' by the other).

Finally, regarding the issue of practicability, probability sampling or recruiting procedures for such a population were treated as impossible, or practically unfeasible, and a chain-link ('snowballing') strategy was adopted.

Recruiting the sample

The first image that we worked with is that of an iceberg: MSM form a group of unknown size and composition, but can be thought of in terms of a 'floating' fraction – the tip of the iceberg – who are 'out', informed gay men, who are also usually more middle-class and better educated – the typical characteristics of volunteer respondents (Rosnow and Rosenthal 1975), who are the easiest and therefore the cheapest financially to contact and persuade to agree to be investigated. They make up the bulk of most studies of gay men. But there is also a submerged fraction, of considerable but unknown size who in various degrees are closeted, unreconciled to their sexual behaviour, possibly leading an apparently heterosexual lifestyle or who are very 'hidden'. These are difficult (and hence expensive) to contact and hard to recruit, and feature rarely in studies of gay men. The need for a 'representative' (let alone a probability) sample is to 'burrow down' into the submerged fraction.

Now the image changes. The crucial thing that makes a population of MSM is that they are not isolates, but interact with other men even if only for sexual purposes. Because of this, MSM form not only a category but a network, which in graph theory terminology is 'connected' – in principle, you can trace a chain of sexual contact from any one to any one other. The sampling procedure which has been developed in recent years for rare, hidden or elusive populations of groups such as MSM or drug-takers is the 'snowball' (or chain-referral or link-tracing) sampling (Goodman 1961, Biernacki and Wald 1981, Spreen 1992). The technique is used to identify some starting set of members of the rare, hidden or hard-to-reach population, ask them to nominate or identify others, who in turn are asked to identify others, and so on.

This tracing procedure was implemented in the Cardiff sample by the 'nibbling' procedure (Rees 1988). What is interesting is that even in large nation-state human populations it does not take more than about seven such links on average to reach any person from any other – the so-called 'small-world' phenomenon (Milgram 1967, Pool and Kochen 1978).

The process used in SIGMA was two-stage: first to obtain easily-accessible respondents in each of the nine project design typology cells (chiefly from gay pubs, clubs and voluntary organizations); second, to use these initial contacts as starting samples for producing tracing trees. In practice the interviewer asked the initial
respondents to name other potential respondents who were of the same (project) type as themselves, but preferably less 'out' as gay. It was left to the interviewer to satisfy him/herself that this definition was understood by the respondent, and we were rarely able to ascertain whether this had actually been done.

The attempt by SIGMA to implement tracing sampling was noble, but ultimately deficient, and for a number of instructive reasons. Often a given gay man's friends and acquaintances are not of the same age-relationship type as himself so that it was frequently quite difficult for a respondent to name someone of the same type, let alone someone who was less 'out'. The number of contacts to be named was never specified; more relevantly, no criterion was provided by which the respondent could decide when the number of his nominees was sufficient. As a project we had bound ourselves to anonymity in the form of not recording or making use of the name of anyone named in the research context. We therefore had to rely upon the respondent to contact his nominee and ask him to participate in the project. Consequently we might never know that a specific person had been thus nominated, let alone whose nominee he was.8

But it must be said that the exercise was not a resounding success, and that the initial SIGMA sample was by no means a fuller 'snowball' than other such studies.

The gay scene in Cardiff and area is a good deal more closely-knit (on any significant criterion) than that of the sister site of London. Where it was possible to track the contacting process in Cardiff, for sexual contacts (of whatever variety) there turned out to be a goodly number of cross-cutting circles, but with weak links between them, so that an estimate of ultimate connectivity probably depends rather importantly on whether the sample includes the liaison persons (bridges) that mediate such clusters. It would also lead to missing certain important subsets of respondents who come in and out of the scene on an occasional basis and who would only normally be contacted via one man; occasional (but not hardened) users of 'cottages' are an important example of this.

This assessment of the topology of the homosexual network is largely impressionistic, and would need to be investigated directly. In neither main site, however, did we normally exceed a chain-length (let alone a tracing step-length) of more than three, and we argued that a length of four would be necessary to achieve even reasonable coverage and more than that would be necessary to come anywhere near encompassing a coherent cluster (Davies 1986).

In the event, then, the correct description of our 'sample' would be that it was a two-factor design quota sample that used multiple site, time and place sampling and snowballing techniques to access respondents. It was planned before the National Sex Survey and it had been hoped that it would be possible to calibrate our structure against their findings and weight appropriately. In the event this has not been possible.

Sexual Diaries

The diary is a chronicle we are all familiar with, even if only as much in its early abandoning as in its keeping. It is extensively used as a form of autobiographical memory, and as Miss Prism goes on to advise Cecily, 'Memory ... is the diary that we all carry about with us'. As a resource for qualitative data for social scientists and historians alike the diary is invaluable; it is a potent 'Document of Life' as Ken Plummer (1983) describes it.

Sexual diaries have been kept quite naturally by all sorts of people and for all sorts of reasons: as a record of prowess, as an information-source in case of infection, as a therapeutic tool. But for whatever reason, gay men have been especially prone to keep such diaries, or to commit to their autobiography the detail of sexual happenings which are diary-like in content and form—perhaps because if society defines gay men by reference to their sexual activity a diary of that sort reaffirms their identity on a daily basis. The sexual diaries of Joe Orton are notorious, but typical of the genre. His editor cites a well-known instance of an encounter with an Irishman shortly after the funeral of his mother:
Friday, 30 December 1966, Leicester. It was an empty house... he didn't live there. He rented it for sex... The bed had springs which creaked. First time I've experienced that. He sucked my cock. Afterwards I fucked him. It was difficult to get in. He had a very tight arse. A Catholic upbringing, I expect. He wanted to fuck me when I'd finished. It seemed unfair to refuse after I'd fucked him. So I let him. We lay in bed and talked for a while... (Lahr 1986, p. 45)

In the course of the SIGMA researches I have come across several instances of men who have kept such a diary for decades (often over their entire sexually active life), but few survive the ravages of hostile families and new one-to-one committed relationships.

In the case of the SIGMA sexual diaries, several of us had kept diaries of this sort for a number of years before the study began, and our construction and systematisation of the diary method grew out of our experience as sexually active gay men, and thus allowed us to tap in immediately to the vocabulary and to the set of distinctions used in everyday discourse among gay men.

Sexual diaries may exist as a social practice and as a resource, but why develop them as a specific method? Basically because of the shortcomings of conventional methods of research when applied to sexual behaviour.

If we were interested primarily in the actual mechanisms of sexual transmission of HIV, then direct observation might well be the appropriate method for obtaining information, as in the Masters and Johnson (1966) study. But such a method is likely to lead to highly biased estimates, since only a highly atypical sub-population is likely to consent, and the presence of an observer would itself be highly reactive. Direct systematic observation as a method is therefore far from unobtrusive and would involve massive problems of consent and organization (and cost). It would in any case be illegal, at least in England and Wales (if not in Scotland) for homosexual activity.

So to obtain relevant information, we normally rely on subjects' own reports or accounts of sexual activity, and the interview setting provides the most usual context of data-collection. In Project SIGMA the yearly core question schedule includes as a central element the "Inventory of Sexual Behaviour" (ISB) (Coxon 1992b), asking respondents a systematic set of questions about whether (and if so, how often) they had engaged in these detailed activities ever (for prevalence) and within a given period of time (for incidence). But how accurate are such subjects' estimates likely to be? From the outset of our enquiries the data gave good grounds for scepticism: the numbers given in answers were often suspiciously vague, rounded or approximate (Coxon 1986b), suggesting problems of accurate memory recall. Moreover, when questions (identical or implied) were repeated later in the interview, the number given was rarely the same, suggesting problems of reliability. When cross-checks were made with estimates given by their partners the numbers were (to varying degrees) often at variance with the respondent's, suggesting problems of validity.

Much information obtained about sexual activity in the interview context is also atomistic and out of context — we learn whether or how often something was done, but rarely the context in which it took place, the sequence in which it occurred or the person with whom it occurred. But as we shall see, such factors make a big difference to the meaning of sexual behaviour, and they are also important in attempting to understand sexual risk. The fact that sexually risky behaviour takes place is important, but if people are to be encouraged to lessen or avoid risk then we need to know the significance of such behaviour to the person, and we also need to identify its context in order to find out whether risk-taking varies systematically by situation, rather than simply by individual. Again, the number of sexual partners a person has is an important variable epidemiologically, but it is even more important to know whether they are one-off or regular partners, whether sex with such partners involves penetrative or unprotected (risky) sex, whether alcohol or drugs such as 'poppers' have been used (possibly as disinhibitors), and so on. In the interview context such questions are asked
separately, and even if recall is excellent we can know nothing about how they co-
occurs with sexual behaviour and combine in a particular sexual situation to increase or
decrease risk. Finally, the order in which
sexual activity occurs (and the position in a
sequence in which an act occurs) can have
quite different effects. An example is the
differing risks of hepatitis infection when
sucking follows fucking as opposed to pre-
ceding it. Similarly, the probabilities of
transmission are very different according to
whether a person is anally receptive or
insertive, and we knew little indeed about
the prevalence and possible mixture of sex-
ual role playing in male-with-male sex.

Advantages and disadvantages

What are the advantages and disadvantages of the sexual diary method? ¹⁰

The sexual diary method is a more ‘nat-
ural’ method than most other methods, both
in the sense that it exists as a common
social practice and that it is written in nat-
ural language. It makes it possible to obtain
information in far greater detail than other
methods, since it is designed to minimize
recall and memory errors and cognitive
strain. It is especially adapted to gathering
reliable information on the time-sequence
of events, so that change is more easily
charted. The information can be obtained
in a contextually-specific manner, without
relying on recall; thus variation due to, e.g.,
particular sorts of partners or particular set-
ings can be studied directly. Quantitative
information is derived directly from the
data, without recourse to the error-full esti-
mating procedures used by survey
questioning or respondent recall. The sex-
ual diary can be augmented to obtain other
concurrent information such as alcohol and
drug use in sex (see Weatherburn et al.
1993), and the data obtained are (on present
evidence) more reliable and valid than
those obtained from retrospective recall in
surveys (see Janson 1990).

These advantages are impressive, but
need to be balanced against the undoubted
disadvantages, some of which can be ame-
liorated. The main disadvantages of the
sexual diary method have to do with bias in
recruitment of respondents rather than
with the method of data collection per se.
But there is undoubted selection bias with
respect to those who do and those who do
not agree to be diary respondents or return
information. In the case of hidden popula-
tions like gay men selection bias in the
recruitment of those prepared to keep a
diary exists in addition to that in the initial
sampling procedure of project sample
members. The sources of bias are very simi-
lar to those in other studies relying on
volunteer subjects (Rosenthal and Rosnow
1975, p. 225): educated, higher social
class, intelligent, approval-motivated and
sociable (the last characteristic here takes
the form of being more likely to be ‘out’ as
gay men). Those volunteering tend to be
more sexually active (in the sense of having
more sexual sessions and more partners)
than those who do not volunteer.¹¹ For lon-
gitudinal studies there is undoubted
‘step-wise attrition’ – it is far easier to per-
suade men to keep a diary for consecutive
months than regularly on a yearly basis.
Finally, the type of data generated cannot
readily be analysed by conventional pack-
ages, and rely on an intermediate stage of
string-manipulation software.

The structure of sexual action

The sexual diary method, like the questions
on sexual behaviour in the core section of
the SIGMA interview schedule (SIGMA
1995), is based upon a common schema of
sexual behaviour, originally developed to
apply to homosexual behaviour, but readily
extendable to other orientations.¹² It is
important to understand the schema before
explaining the diary method further. A dia-
grammatic version of the schema is
presented in Figure 2.1.

The unit of sexual behaviour is the Sexual
Session, denoted by the outer square box in
Figure 2.1. Using a linguistic analogy,
which we shall later exploit more literally,
the sexual session may be considered as the
‘sentence’ of sexual activity – self-sufficient
and intrinsically well-formed.

The sexual session occurs at a given
(specified) time and place and is made up
of one or more sexual acts (in the rounded
Each sexual session involves at least one, usually two (and sometimes more) people. It is typically ended by sleep, a non-sexual intermission or change of partner. The characteristics of a Sexual Session thus include four components: the setting that contextualizes the sexual act, the antecedents, the accompaniments, and the partner specification.

The setting refers primarily to where and when the sexual activity took place. 'When' is usually the time of day, so that sexual sessions can be kept in time order, and the 'Where' usually refers to the location (such as 'boyfriend's flat' or 'my home', or outside locations such as 'public toilet at X', or 'Park Y').

The antecedents refer to any relevant events which preceded or led up to the sex, such as prelatory drinks or drugs, use of poppers (nitrites) or stimuli such as videos or magazines.

The accompaniments refer to antecedents which continued during the sexual activity, together with the use of additional objects during the sexual activity, such as lubricants and 'toys' (e.g. dildoes, nipple clamps, leather).

The partner specification would ideally consist of both the name and characteristics of the partner/s, such as age, sex, relationship status. However, it is often necessary to remove the actual name of a partner from the record when undertakings of anonymity or confidentiality have been given, but in so doing all record linkage (for network or contact analysis) is then forgone.

The core component in describing sexual activity, the 'atom' of sexual behaviour, is what actually happens - in our terminology, the 'sexual act'. On the linguistic analogy, the sexual act is the word in the sentence, and sexual act(s) make up the sexual session. It is the sexual act which specifies 'WHO does WHAT and with what EFFECT'. These three components of the sexual act are referred to as: the behaviour (or behaviours) which refers to the actual sexual activity itself ('what' is done); the modality which refers to 'who (does the activity), and to whom'; and the outcome which refers to the 'effect' of this sexual activity, which in the context of HIV transmission becomes the question of whether ejaculation of sperm occurs, and if so, who ejaculates, and in what manner. Let us specify each of these in turn.

**Behaviours of the sexual act**

A succession of what are actually continuous bodily movements are, or can be, 'chunked' into identifiable and (well-nigh) universally recognizable sexual activity and given a common name. For instance, whilst the act of masturbation will usually have at least some unique components (for no one does it exactly the same way, and no one repeats the act identically), the stimulation of the penis by the hand is usually taken to be a necessary part of the definition. As in so many other aspects of sexual behaviour, even such a behaviour as Kinsey insists that it also has an intentional aspect, so that random, unintended rubbing of the penis would be thus excluded from the category of 'masturbation'. Despite the fact that virtually every man has done it and that it is done from an early age, the nomenclature is far from universal. Because the act of masturbation is often taboo especially among children, and frequently discouraged or prevented, it comes to be referred to by all sorts of euphemisms and code-names (often unique to the family concerned). There thus arises a hierarchy of terms of differing acceptability, from the medical terminology used professionally (and often when talking to professionals) through a widely-used set of vernacular
terms to largely idiosyncratic ones. In Project SIGMA, the WHO studies and in related studies we have always elicited such terminology before proceeding to questioning on detail of sexual behaviour (see Appendix 5). The purpose of this is not only to gather information on ‘street’ terminology, but also to make the respondent more at ease in asking detailed information about what may be an embarrassing topic. For this reason, all our research forms including the Inventory of Sexual Behaviour are constructed with the neutral (medical) term in curly brackets, and the subject’s chosen alternative is substituted in the spoken form. Thus: ‘Have you ever [masturbated] another man to his [ejaculation]?’ will be rendered in the interview as: ‘Have you ever “wanked” a guy until he “came”? if those are the subject’s preferred terms, or indeed as: ‘Have you ever “given a hand-job to” a guy until he “spunked”? in another case.

The set of sexual behaviours must be open-ended. Although the number of distinct and distinguishable sexual activities is immense, the distribution of types of sexual behaviours is very skewed and a surprisingly small handful of behaviours (typically three) normally suffices to encompass over 95 per cent of sexual behaviours.

It is not usually difficult to specify the list of the more common or well-established sex behaviours such as wanking and sucking; it is the more recently developed and ill-defined ones which cause difficulty. Thus ‘massage’ can mean anything from rubbing hands over the body to systematic manipulation of muscles, and a practice like inter-femoral penile insertion (‘thigh-fucking’), which tends to substitute for anal intercourse in a post-AIDS era, can cover a wide range of actual practices. In the case of project SIGMA and the WHO Homosexual Response Studies, the list of sexual behaviours given in Appendix 1 is used (gay men’s sexual behaviour is discussed in detail in Chapter four).

Modality of the Sexual Act
The modality of the sexual act – the prefix to the sexual act ‘word’ – specifies which actor (from ego’s standpoint) did what sexual act to which actor. A variety of contrast terms exist to distinguish these two roles: inserter/insertee, active/passive, as well as ‘street’ language such as the ‘doer’ (usually, inserter), butch/bitch, etc. Although particularly relevant to anal and vaginal intercourse, the inserter/ee contrast cannot be used for all sexual behaviour since it carries with it a view of sexual activity as primarily insertive (and by implication, male). More importantly, the role-difference in some behaviours – such as wanking – is not, and cannot reasonably be described as, that of inserter/ee. Therefore we adopt the ‘active/passive’ distinction as basic, and use it in accord with conventional grammatical usage to denote verbal mood (hence ‘modality’), noting that this also conforms to common usage among gay men:

‘Active’ means that ego does the sexual act to alter.
‘Passive’ means that the sexual act is done to ego by alter.

Thus, ‘active masturbation’ means that I masturbate the other person and ‘passive masturbation’ means that the other person masturbates me. Besides the basic active/passive modalities, we further distinguish:

‘Mutual’ means that ego does the sexual act to alter at the same time as alter does the identical act to ego.
‘Self’ means that ego does the sexual act to himself.
‘Him’ means that alter does the sexual act to himself.

These modalities are discussed further in Chapter five.

The outcome of the Sexual Act: Ejaculation
The dominance of HIV transmission in our account naturally means that we concentrate primarily on ejaculation of sperm (sometimes this is inaccurately but euphemistically called ‘orgasm’) as the primary ‘outcome’ of the sexual act. Initially we concentrated simply on whether a subject ejaculated or did not as the result of a given sexual act. As time went on the advice of safer-sex guidelines (and the inventiveness of some members of the gay
community) meant that we had to pay more careful attention to exactly where the ejaculate went, and by the increasing need to know whether that destination was a condom. Thus condoms began as ‘accompaniments’ in our schema and finished as an integral part of the outcome.

After an initial attempt to work with a simpler description of ejaculation (see Coxon et al. 1992, pp. 68–9) it became necessary to allow five possible outcomes for each of the partners: into a person, onto a person, into a condom, some other destination and no ejaculation. These are described and elaborated further in Chapter five.

Before putting all these components together to lay out the structure of sexual behaviour, three further issues must be briefly discussed: the representation of complex sexual behaviours, the representation of their sequencing, and the description of the accompaniments of sexual activity.

On occasion two (or sometimes more) sexual acts occur simultaneously and in a way that may have implications for HIV transmission. For instance, wanking is often combined with fucking, but in very different forms. In this case we introduce concatenation as an operator to combine two acts done simultaneously and symbolize it by an ampersand, ‘&’. Examples commonly include:

AF&HW (Active Fuck & He Wank: i.e. Ego fucks Alter anally and at the same time Alter is wanking himself)

PS&ATt (Passive Suck & Active Tit-torture: i.e. Ego is being sucked and at the same time is tweaking Alter’s nipples.

Such complex (concatenated, simultaneous) sexual acts are treated as a single act. If a sexual session contains more than one sexual act, some conventions are necessary to denote the sequence in which they occur. This is done simply by linear sequencing, using a space as a separator between acts:

AW PW AS PS MW, reading in longhand as: Active wank, then Passive wank, then Active suck, then Passive suck, then Mutual (simultaneous) wank.

The antecedents and accompaniments give information about the context or setting of the sexual behaviour which might have relevance for HIV transmission. Although it is useful to distinguish antecedents and accompaniments in this way it has a degree of artificiality since there will be some things which continue both before and throughout a session, others which are used intermittently (e.g. ‘poppers’ or bondage), and yet others which will refer only to a specific act (e.g. the use of a lubricant or dildo or a condom). Those which are of special relevance are included in the table below.

These are discussed further in Chapter five.

**Instruments and formats**

The sexual diary method has changed little since its development and inception in

<table>
<thead>
<tr>
<th>Class</th>
<th>Examples</th>
<th>Sexual Behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Drugs</td>
<td>amy/ nitrates, cannabis, ecstasy etc.</td>
<td>especially fucking</td>
</tr>
<tr>
<td></td>
<td>alcohol</td>
<td></td>
</tr>
<tr>
<td>2. Prophylactics</td>
<td>finger stools, dental dams, veterinary</td>
<td>fingering</td>
</tr>
<tr>
<td></td>
<td>gloves</td>
<td>rimming</td>
</tr>
<tr>
<td>3. Lubricants</td>
<td>water based (e.g. ky)</td>
<td>fisting</td>
</tr>
<tr>
<td></td>
<td>oil-based</td>
<td>especially fucking</td>
</tr>
<tr>
<td></td>
<td>vegetable-oil based</td>
<td></td>
</tr>
<tr>
<td></td>
<td>saliva, semen</td>
<td></td>
</tr>
<tr>
<td>4. ‘Toys’</td>
<td>restrictors (‘cock-rings’)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>’pain-pleasure’ (belts, nipple-clamps, pendant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>weights)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>stimulators (dildoes)</td>
<td></td>
</tr>
</tbody>
</table>
Men, Diaries and Other Techniques 25

1983. Although a retrospective one-week version was written into the SIGMA core questionnaire from the outset, the normal format consists of a month-long diary which is self-administered, and hence instructions are needed which will be readily understood. The current instructions are reproduced in Appendix 2.

The primary purposes of the instructions are to alert the potential diarist to the characteristics of the sexual behaviour schema outlined earlier and stress the importance of keeping as small a gap as possible between the events and their record. In the current form, the diarist is encouraged to write in natural (if telegraphic) language — in the past they were encouraged to write in code, at least in part.17

The other components of the Diary Pack are reproduced in Appendix 3 and consist of (1) face-sheet information (2) partner list and (3) example of a one-week sheet of the sexual diary. Diarists are encouraged to spread over as much space as they wish, and some dispensing entirely with the formatted sheet and write their diary in free format.

To give the feel of diary-data, an anonymized (but close-to-original) week of a sexual diary is reproduced in Figure 2.2. This particular account is very factual and brief and the entries obey most of the guidelines. All ejaculations are unambiguous in terms of who ejaculated and whether it was protected, if not as to the exact destination of the ejaculate. Very little ‘taken-for-granted’ pragmatic information is needed to interpret the diary entries, with the possible exception of ‘69’, mutual/simultaneous sucking. But such simplicity is not always a feature. Consider the following (genuine) entry:

FRI 5pm @ P2’s flat after meals and bottle of sherry between us both on poppers fucked each other and wanked each other after an hours break.

10pm I sucked him (P2) hard ‘till he came in my mouth & I swallowed all his cum

In the 5pm session, the men ‘fuck each other’: clearly they did not, nor could they, sexual coder assumed), so the fucking must have been sequential, but we have no clear indication of who fucked whom first. Did either come? If so, when fucking or when wanking? It is not entirely clear, but pragmatic information about gay sex and contextual information about this man’s typical practice would indicate that the ‘hours break’ may have been to regain momentum after coming as a result of (active) fucking, and that they each came a second time as a result of the (simultaneous?) mutual wank.

Encoding and Representation
And so we come to the more esoteric parts of sexual diary data: their representation and encoding. The schema of sexual behaviour described above can be given a more formal logical representation, outlined in Appendix 4. Once well-defined, the components and structure of the schema are readily mapped onto the structure of the database record used to store diary information.

The core of the sexual schema is the sexual act, and it is here that the encoding is most specific; in other fields the entries can simply be extracted from the natural-language accounts. Each sexual act consists of a behaviour done by the actor/s in a specific way (modality) and possibly resulting in one or both ejaculating. Each act is therefore given the form:

M modality
B behaviour
E ego’s ejaculation
A alter’s ejaculation

Thus the ‘root’ of the act is the behaviour, the modality is the prefix and the outcome/s are the suffix/es. When encoding a given act, each alternative is represented by a letter (though some behaviours are represented by two letters) which are chosen to be as close as possible to the natural language ‘street’ term. These abbreviations are as follows:

<table>
<thead>
<tr>
<th>Modalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>E</td>
</tr>
<tr>
<td>A</td>
</tr>
</tbody>
</table>

S Self
# SIGMA SEXUAL DIARY FORM

**ID Number:** _____/_____/5  
**WEEK BEGINNING:** 3/MAY/1992  
(day) (month) (year)

---

**Remember, each session should include:**  
- The Time, The Place, The Partners (from partner list)  
- Then, the session in your own words  
- If you "come" (ejaculate) in the session, remember to be explicit about where it goes and **always** to record the use of condoms.  
- List any accompaniments you use (poppers, lubricants, drugs, sex toys, ...)

| Sunday | 9am My flat, P1  
  | We deep kissed, and moved into a "69". Whilst doing it I began to finger him. Then he wanked me (both using poppers) and I came. Following that I wanked him till he came.  
  |  
| Monday | 12.30pm Lunch-time wank at work; I didn't come  
  |  
| Tuesday | 11.30pm, Hampstead Heath, P2: We wanked each other off; both came. 12.15am (Thurs) P3; I sucked him, then he put on a condom and fucked me; he came.  
  |  
| Wednesday | (see above: Wed)  
  |  
| Thursday | 8.30 reading porn: quick wank to orgasm.  
  |  
| Friday | 11.00 p.m. P1's flat. After eve at The Bell (4 pints), I sucked P2, then he fucked me and came (no condom).  
  |  
| Saturday | 7.30 a.m. I woke up to find P1 wanking me. Then he sucked me off, and I came in his mouth. We began using poppers and I sucked him, carrying on to fuck him (with condom), whilst he wanked himself. He came, I didn't.  
  |
There are four acts in this session (marked off by the # lines). Each is taken separately and put into the act-form (see table below). So the entire session is encoded as: (MDK MS & AFg PW, XN /p AW, NX).

The finer details of encoding sexual diary information and entering the data are contained in the Instructions for Encoding Sexual Diaries, which forms the Coders' Manual used by diary encoders.18

Each session is then entered as record in the database (currently using Cardbox-Plus™ and dBaseIV™ formats). An example, encoding the illustration of the first session in the one-week diary, is as follows:

```
NO: XXX | TYPE: IV | STATUS: NEG | DATE: 03/05/92
TIME: 09:00 | PLACE: OWN FLAT

PARTNER: F1

ACT: MDK MS & AFg PW, XN /p AW, XN
POPPERS: Y | CONDOMS: N | LUBS: N

OTHER:

DRUGS: N

DIARY START: 03/05/92 | DIARY END: 31/05/92 | RECORD CODE: S
```

Once entered in the database any conventional selection, counting and logical retrieval can be done on the diary data. But at an early stage it was found necessary to produce special-purpose software for the analysis of the data, which are a curious amalgam of linguistic, verbal and quantitative information and call for a unique combination of methods and styles of analysis. The earliest programmes were written by Mr Chris Mitchell (of the eponymous 'Chriscode') and the suite were re-written, extended and systematized into a suite of C++ programs called SDA (Sexual
Data-sets

Data-sets used in this and subsequent chapters span eight years of the AIDS pandemic, from 1986 to 1995, and include both ‘one-off’ diary surveys and diaries from the longitudinal SIGMA cohort.

GT86 (1986; N = 188)
The *Gay Times* data-set (‘GT86’) originates from a feature article by the author titled ‘Gay Sexual Lifestyles’ in the May 1986 edition of *Gay Times*. The article announced the start of funding from the Medical Research Council for Project SIGMA, but I had been developing the diary method over a number of years (the first published diary data occur in Coxon 1986 and refer to a diary collected in 1984): 273 men responded to the *Gay Times* appeal to keep a sexual diary for a month; 188 diaries (69 per cent) were received back. Compared to the general population of males, respondents over-represent the South-East and under-represent Scotland and the North-West. They were strongly over-representative of the 21–39 age group and strongly under-representative of the 40+ age-group. The typology percentages for GT86 are shown below.

**| Relationship | Age | Under 21 | 21–39 | Over 39 |
---|-------------|------|---------|--------|--------|
| Closed      | (I)  | 1.8%   | 12.9%  | 4.1%   |
| Open        | (IV) | 2.3%   | 20.5%  | 14.6%  |
| No regular  | (VII)| 5.8%   | 28.1%  | 9.9%   |

GT86 \(N = 171\) (17 n.c.)

SIGMA (1986–93; N = 385)
Each year, after the interview, SIGMA panel members were asked to keep a sexual diary for a month. Their diaries form the base of the main SIGMA data-set and represent Waves 1 to 5 of data collection, and thus span 1986 to 1993. The 385 men came primarily from London and South Wales, with others from Newcastle, Teesside, Birmingham, Portsmouth, Leeds, Norwich, Liverpool and Bristol. Relatively few respondents kept a diary in each wave, and therefore only a fraction of these data are fully longitudinal. This forms the largest single data-set of sexual diary data. The typology percentages are shown in the lower table below.

There is quite a strong similarity between GT86 and SIGMA in terms of the percentage distribution over the age-by-relationship table; given the totally different methods of recruitment, the similarity is striking.

| Relationship | Age | Under 21 | 21–39 | Over 39 |
---|------|---------|--------|--------|
| Closed | (I)  | 1.6%   | 11.7%  | 4.0%   |
| Open  | (IV) | 1.1%   | 25.2%  | 18.5%  |
| No regular | (VII)| 9.5%   | 15.6%  | 12.7%  |
Men, Diaries and Other Techniques 29

GP93 (1992–93; N=79)
In November/December 1992 an appeal for men to keep a month sexual diary was made in the gay press (hence ‘GP93’ data): Boyz, Pink Paper, Capital Gay (weeklies) and Gay Times (monthly). With the exception of Boyz, the response was disappointing. A total of 178 men wrote in for a diary and 86 returned a completed diary, of which 79 were usable: 60 per cent from Boyz, 22 per cent from Capital Gay and 17 per cent from Gay Times. A significant fraction of these men were persuaded to keep a diary for one or more – in a few cases, five – subsequent months, so that longer-term stability of patterns of sexual behaviour could be investigated. The 1992/3 study was also a study of validity and reliability of the diary method. Having completed and returned a diary, respondents were then written to and asked to estimate the number of times they had done various sexual activities in the period of the diary, thus simulating a ‘questionnaire’ or ‘interview’ response which could be compared directly to the actual count of their activities from their diaries. The typology percentages for GP93 are set out below.

This distribution is rather different to the previous two, though the same structure is clearly present. In particular, open relationships are fewer and casual (no regular) relationships have increased. The original diarists of GP93 were repeatedly invited to keep further diaries (in some cases up to six such repeat diaries exist). In Chapter eight, an extended version of the original set is analysed, taking data up to November 1995.

Negotiations are proceeding to lodge the SIGMA diary data and programs in the ESRC Data Archive at the University of Essex, to complement the SIGMA interview data already held there and the (anonymized) data will be accessible to any bona fide researcher. The following statistics of the diary data hold as of 1 January 1996.2

53,877 sexual acts in
25,082 sexual sessions, performed by
1,035 individuals/months from
10 locations/sites.

Diaries and interviews: validity issues
Under funding from the Department of Health a number of studies have been done on the reliability and validity of these diary studies of sexual behaviour. These have included: comparing the data on sexual behaviour derived from SIGMA interviews and the data from their subsequent month diaries for a subset of SIGMA panel members (usually adjacent months). Gay men who made up the GP93 diary data-set were asked after returning their diaries both to estimate the frequency with which they had done various sexual acts in the diary and rate how sure they were of their estimates. (The subsequent estimation procedure was done to simulate an interview situation, but referring to the same month, thus forming a unique comparative data set.) The main findings are reported in Coxon 1995a. The ‘headline’ results are as follows:

86 per cent of the diarists returned an estimate form.

The more frequent a sexual behaviour, the less certain diarists were of their estimate. Correlatively, the rare acts (especially fucking) have higher certainty ratings.

The profile correlation for sexual behaviour averages between ‘interviews’ and diary-counts is ρ = 0.968 (linear) and τ = 0.655 (ordinal), but ‘interview’ estimates of behaviour are (on average) consistently higher than counts made from their diaries.

This holds also on the individual level; two-thirds of acts are over-estimated in the ‘interview’ compared to the diary-counts and only one-fifth of the sample have identical ‘interview’ estimates and diary-counts.

---

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Age-under 21</th>
<th>21-39</th>
<th>Over 39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closed</td>
<td>(I) 6.3%</td>
<td>(II) 6.3%</td>
<td>(III) 2.5%</td>
</tr>
<tr>
<td>Open</td>
<td>(IV) 3.8%</td>
<td>(V) 19.0%</td>
<td>(VI) 6.3%</td>
</tr>
<tr>
<td>No regular</td>
<td>(VII) 5.1%</td>
<td>(VIII) 30.4%</td>
<td>(IX) 20.3%</td>
</tr>
</tbody>
</table>

GP93  N = 79
The major individual discrepancies (inaccuracies) between 'interview' and diary-counts are located in 'fucking without condom'.

Work is currently proceeding on estimating 'selection bias' among diary-keepers (i.e. to what extent do diary-keepers differ from non-diary-keepers in their sexual behaviour?), alternative methods of collecting sexual diary information (e.g. Internet response), and feasibility studies in extending diary methods to heterosexual and minority ethnic group men and women.

**Supplementary methods**

In practice, the project questionnaire and sexual diary data collection were closely intertwined, though a number of diary respondents were recruited quite independently. But relatively little use is made in this volume of the interview data and it does not therefore need separate discussion. On the other hand, the Cardiff site in particular made use of other methods such as participation and/or observation, and this material is reported here.

The material in Chapter seven on 'Different Scenes' is primarily based upon participation and/or observation. The investigators and most staff members were mostly social-science trained and themselves gay, so participation was never a problem in most of the areas referred to in this book (this is not to say that gay males were necessarily better as researchers). Systematic observation and/or participation is another matter and since the topic areas chosen are minority pursuits (basically, cottaging and leather/SM) and strongly dependent on individual penchant, observation or participant involvement in these scenes was never required of staff.

**Cottaging and cruising**

The cottaging studies were carried out in the period 1982–88 in South Wales and in 1988–95 in Essex by the Cardiff and Essex sites²³ of Project SIGMA. Those involved in the studies included the principal investigator, part-time research staff of the project who were recruited specifically to undertake this work, and those permanent members who chose to be involved (it was not part of their job-description), together with SIGMA panel members who were known to be active cottagers and who agreed to be thus involved.

Since cottaging and cruising can be construed as an illegal activity (under various legal headings including Vagrancy Acts and more recently under the Criminal Justice Acts) some accommodation had to be reached with the police and other authorities, since employees of the project could not be expected to engage in what could be construed as illegal activity. Initial approaches to the South Wales Constabulary met with total opposition, chiefly on the grounds that they could not provide protection for researchers against violence which they maintained the researchers would be subject to. After some negotiation it was suggested that if we wished to persist with the studies this would be acceptable so long as no offence was committed and so long as Police authorities were informed of where we were working. The latter condition was clearly unacceptable, and a final uneasy compromise was reached whereby researchers carried an authorizing card with photo which they would produce in the event of a police raid and arrest. In Colchester the police authorities were more co-operative and a set of arrangements and understandings were worked out by which police were informed of the identity of researchers and their car numbers, and researchers carried the ID cards. Researchers were advised not to present the card in the case of arrest until being charged.

The strategy used in covering likely sites was threefold: to carry out a 'census' of toilets and known cruising areas (in the greater Cardiff area, in the South Wales valleys between Newport and Swansea, in central Bristol and in Liverpool and Manchester sites) to determine whether there was any sign of homosexual activity; to restrict systematic investigation to a few regular 'hot' sites and use time-sampling techniques to cover various periods; and to visit other sites intermittently. A similar strategy was used in Colchester and
surroundings, but covering a smaller area.

A ‘cottage kit’ was developed to facilitate systematic description and report. The ‘Field Research Site Record’ records the site, plan of location, internal plan and description (typical clientele, active periods, police surveillance and warnings, graffiti content and dates). The ‘Field Research Observation Record’ is filled out for every visit, indicating: time, date and weather conditions; personnel (cottagers): age, build, attractiveness, sexual show; clothing: jeans, leather, suit; keys, handkerchiefs and other semantics; vehicle number (and apparent origin); sexual activity: timing (period of time spent in cottage or cubicle, escape of time before others’ moves); movements, termination behaviour (talking, exchange of notes, tryst arrangement).

Leather/SM

The leather/SM research depends on two studies: the Hunt study and the Coxon study of somewhat different form.

Hunt study

In the second wave (1989) of interviews of Project SIGMA a set of questions (D7; D7.1 to 7.6) were included which related to SM, concluding with one which said: ‘Would you be prepared to talk about your SM experience in more detail to another researcher?’ The questions were analysed (and further interviews conducted) by Andrew Hunt, then Senior Researcher at the London SIGMA site and himself a participant in the leather/rubber/SM scene. The 691 respondents in Wave 2 who answered these questions thus form a unique, large and fairly representative sample of gay men and their attitudes to SM, as well as providing a sub-sample of those involved in the area. His report remains unpublished except as a working paper (Hunt 1991) and the data reported here for the first time are drawn from that report.

Coxon study

Townsend’s Leatherman’s Handbook contains a simple questionnaire and some summary results from his own studies. During the mid 1980s this was modified and extended by the author for his own personal use as a participant on the leather/SM scene, and later systematized (Coxon 1991) and used more extensively by being given primarily to other masters for distribution to their slaves as part of the scene. This networking procedure has produced 135 completed questionnaires and has clearly penetrated the SM scene to a considerable degree, primarily in London and the Home Counties, but also in the Midlands, North of England and Scotland. This work has been augmented by the post-graduate research work of Wouter Geurtsen (University of Tilburg) on the leather scene in London and Amsterdam and supervised by me in his English work.

The respondent sample is clearly biased towards slaves, and to those undertaking longer-term commitment rather than one-night encounters. It also under-represents those who do not operate within the ‘SSC’ (Safe, Sane, Consensual) framework. Despite these biases, the data are very reliable and have usually been checked out behaviourally by the other SM partners (usually the masters involved and in some cases also by the slaves involved).

List of practices (behaviours) included in the questionnaire include:

<table>
<thead>
<tr>
<th>Practice</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Cocksucking</td>
<td>(AS)</td>
</tr>
<tr>
<td>Passive Cocksucking</td>
<td>(PS)</td>
</tr>
<tr>
<td>Active Arse-fucking</td>
<td>(AF)</td>
</tr>
<tr>
<td>Passive Arse-fucking</td>
<td>(PF)</td>
</tr>
<tr>
<td>Active Rimming</td>
<td>(ARI)</td>
</tr>
<tr>
<td>Passive Rimming</td>
<td>(PRI)</td>
</tr>
<tr>
<td>Bondage</td>
<td>(BO)</td>
</tr>
<tr>
<td>Whipping/belting</td>
<td>(CP1)</td>
</tr>
<tr>
<td>Cane/crop</td>
<td>(CP2)</td>
</tr>
<tr>
<td>Cat o’ 9-tails</td>
<td>(CP3)</td>
</tr>
<tr>
<td>Fistfucking</td>
<td>(FI)</td>
</tr>
<tr>
<td>Water-sports</td>
<td>(WS)</td>
</tr>
<tr>
<td>Scat</td>
<td>(SC)</td>
</tr>
<tr>
<td>Cock &amp; Ball work/torture</td>
<td>(CBT)</td>
</tr>
<tr>
<td>Tit work/torture</td>
<td>(TT)</td>
</tr>
<tr>
<td>Piercing</td>
<td>(PI)</td>
</tr>
<tr>
<td>Using dildoes/butplugs</td>
<td>(FD)</td>
</tr>
<tr>
<td>Using hood/blindfold</td>
<td>(lh)</td>
</tr>
<tr>
<td>Bootlicking/tieing</td>
<td>(BL)</td>
</tr>
<tr>
<td>Genital shaving</td>
<td>(SH)</td>
</tr>
</tbody>
</table>
Full body shaving (BSH)
Candle-wax (WX)
Electricals (EL)

The fantasy items referred to in the text of Chapter seven apply to the question 'My wank fantasies include . . . ' and are:

Leather clothing
Being stripped
Being bound
Being whipped
Fucking another
Getting raped
Being pissed on
Being humiliated
Being enslaved
Being castrated
Being hanged
Interrogation
Dog scenes
Piercing scenes
Outdoor scenes
Slings/suspension
Stripping another
Binding someone else
Whipping another (guy)
Being fucked
Raping another
Pissing on another
Humiliating another
Enslaving another
Castrating another
Lynching another
Military scenes
Flogging scenes
Group scenes
Other/s (specify)

Blood and saliva testing

Following the extended discussions with the MRC about blood-testing for HIV-1 antibodies in the early stages of the project, all SIGMA respondents were asked to give a blood sample for testing. In the full course of the project over 2,000 blood samples were taken, without a single instance of needle-stick or similar injury. The proportion agreeing to give a blood sample was usually about two-thirds and increased systematically as the project proceeded. Blood samples were taken by interviewers (who had been trained as phlebotomists) after the subject had given informed consent. Respondents could opt to be told or not to be told the result. If respondents wished to know the result of their blood test, the person taking the blood pre-counselling them. If they chose to be told their result they were post-test counselled and given their result by the clinical investigator in the London site (Dr T. J. McManus) or the principal investigator in the South Wales site (the author).

Saliva tests, which had been developed as an alternative to blood-testing (Johnson, Parry, Best et al. 1988), were also used, initially to parallel the blood tests and in a pioneer study to assess the feasibility of using saliva-testing in a field setting (Hunt, Connell, Christofinis et al. 1993). Additionally, saliva tests were administered if the respondent agreed so that the results could be calibrated against their blood result and thus assess the saliva test's specificity (error in falsely identifying a negative result as positive) and sensitivity (error in failing to detect a positive result) to HIV-1. Blood samples were tested for HIV-1 antibodies and for a range of other viral markers, including Hepatitis-B.²⁴

As the panel continued, the original sample became more dispersed geographically and it became less feasible to visit panel members twice – to interview and bleed them, and then to return to give them their result. In later waves it was decided to rely solely on saliva tests (and therefore not give results). In recent times non-Panel diarists are asked to self-administer the saliva test and return it with their completed diary.²⁵

Notes

1. The quotation pre-dates current concerns for non-sexist language.
2. Probably the most influential and important US publication on AIDS behavioural methodology is Ostrow and Kessler (1993). In the Introduction and ten contributed chapters there occur only twelve references to British work or researchers (six of these are to the biologist Sir Roy Anderson), and six chapters have no reference to British work or workers. By contrast, the corresponding British collection (Boulton 1994) has almost
two-thirds of its references to US scholars and researchers... US scholars are good, but not that good.

3. Since the SIGMA study was restricted entirely to males, male terminology is used descriptively throughout the book. Where the reference is to a wider group, care is taken to vary the male pronoun.

4. From asking SIGMA panel members before the UK national study we know that a high fraction of gay men would not agree to be interviewed in the UK National Survey (though if they did, they would largely tell the truth), so real refusal rates of 31 per cent are likely to contain a high number of male homosexuals. Truth-telling in the interview is also questionable, and rates of those admitting to homosexual activity are affected by factors such as the presence of others in the interview. The male homosexual incidence rates reported (Johnson and Wadsworth 1994, pp. 188–93) vary between 6 per cent (any homosexual contact) and 1 per cent (genital contact in the last year) are consequently probably considerable under-estimates.

5. In retrospect it would have been better to have divided the large middle age-category and make allowance for the changed age of homosexual consent to eighteen, making: under 18; 18-25; 25 to 39; over 39 as the categories.

6. See Kalton (1993) and Harry 1990 for good discussions of the feasibility of probability sampling of male homosexuals.

7. 'The world is composed of networks not groups' (Wellman 1988, p. 37).

8. At a later stage we decided that the anonymity undertaking might be a case of shooting ourselves in the methodological foot (see the discussion in Coxon 1993), but there were (and are) excellent reasons why gay men need to be persuaded that such information is safe and cannot be used against them.

9. The presence of a third person (here the social scientist) would render homosexual activity illegal under the 1967 Sexual Offences Act as being no longer a 'private' activity. Scottish law does not contain this restriction.

10. This section relies on Coxon 1994.

11. We have been able to show that it is selection bias toward those with busier sex lives rather than exaggeration of activity that is primarily producing bias.

12. The schema is reported in several places, but the most extensive account is in Coxon et al. 1992, on which this section is based. The applicability of the schema and of the diary method reported here is currently being investigated on bisexual, heterosexual and lesbian subjects.

13. The characteristics requested depend on the purpose of the analysis. In some cases it is feasible to ask the identity, or at least the initials, of the partner. Where this is not feasible (or because an undertaking of strict anonymity has been given, as in earlier stages of Project SIGMA) each partner is assigned a unique (arbitrary, but sequential) number. Partners are described in terms of gender (male, female), partner status (regular, occasional, casual, one-off), age, how long the subject has been having sex with the partner, where the partner was met on this occasion, HIV antibody status (if known), and initials or name (see Appendix 3).

15. Respondents are informed that initials are useful to ensure that the same partner can be identified across different month-diaries and also to help link data when possible. An experiment by the author and Dr Chris Joyce of CSCD Colindale in inferring identity by 'blind' matching of partners using profile-matching and genetic algorithm techniques is reported in Coxon 1995b.

15. 'The word “masturbation” refers to any self-stimulation which is deliberate and designed (sic) to effect erotic arousal. By such a definition the accidental touching of oneself is not masturbation because it is not deliberate' (see Kinsey 1948, pp. 497–8).

16. Indeed, condoms can still be used as accompaniments, as when they are used as a toy, or when they are used as a prophylactic before 'pulling out'. This allows us to distinguish using a condom when fucking (but not coming) from coming into a condom when fucking.

17. This was before the introduction of the complexities of the ‘Chriscode’ (see Chapter five) for recording ejaculation, and it allowed the contents of the diary to be kept from prying eyes, usually of a partner, parent or landlady. But increasing complexity meant also increasing error, so self-encoding was no longer encouraged, with the single exception of ‘SWO’, which in the older system meant ‘self-wank to orgasm’. It also occurred on a Project T-shirt which said ‘Why SWO when you can be interviewed by Project SIGMA?’

18. This booklet is available from Project SIGMA, University of Essex.

19. The program suite (and extensive test data) is available at cost from: Project SIGMA, University of Essex.
20. This study is documented in Coxon 1986.
21. 1 per cent did a diary in all five waves; 6 per cent in four; 24 per cent in three; 62 per cent in two; 7 per cent in one wave.
22. These numbers exclude the original GT86 data and more recent bisexual diaries.
23. Other investigations centred on Bristol, Liverpool and Manchester but only in Liverpool were they carried out for more than restricted periods. I am grateful to John Gay and Andrew Greenlees for their painstaking work in these sites.
24. Blood samples were ultimately tested for a range of viral markers including HIV-1, CMV, HBSANTIGEN, HBVANTICOR, ANTI-HBS, HAVG, HAVM, HCV, VZV, EBV, HTLV-1, HHV6, TOXO, ADENO, DELTA, HSV.
25. Sero-prevalence results from SIGMA are reported in Hunt, Christofinis, Coxon et al. (1990) and Hunt, Davies, McManus et al. (1992).
Parallel accounts? Discrepancies between self-report (diary) and recall (questionnaire) measures of the same sexual behaviour

A. P. M. Coxon
Principal Investigator, Project SIGMA, University of Essex, Professor, Health and Social Services Institute, University of Essex, Colchester, UK

Abstract Questionnaires and diaries have complementary biases and advantages for obtaining information on sexual behaviour but self-completed sexual diaries have the advantage of reducing retrospective bias. In a validation study of homosexual behaviour, sexual diary counts and subsequent questionnaire estimates (together with ratings of the certainty of the estimates) referring to the same month are compared and the discrepancies analyzed. Main findings include: questionnaire data yield consistently higher average estimates than diary counts, but have the same ordinal profile; individual difference (diary-questionnaire) scores show that 55% of questionnaire estimates of acts are higher than diary counts, 20% are identical and 25% are under-estimates; discrepancies are differentially located in different sexual acts. Masturbation and fellatio are systematically over-estimated in questionnaires and anal intercourse without a condom is the major source of inaccuracies, but in different directions: active partners under-estimate and passive partners over-estimate the amount of highest-risk sex. A strategy of joint use is discussed.

Introduction
In recent overviews of AIDS and sexuality research, the measurement of sexual behaviour is still seen as the foremost methodological problem confronting not only research on AIDS, but on sexually transmitted infections generally (Boulton, 1994; Catania et al., 1990; 1993; Fortenberry, et al., 1997; Ostrow & Kessler, 1993; Turner et al., 1997). For reasons of practicality and cost, it is inevitable that survey methods form the main mechanism for obtaining reports on sexual behaviour, though the quality, accuracy and validity of the resulting survey data for this purpose have been subject to considerable questioning (see Miller et al., 1990; Johnson et al., 1994, for the USA and Britain, respectively). The main grounds for disputing the accuracy of survey-based estimates of sexual activity are the reactivity of the interview situation (due both to the impact of interviewer bias and to the socially stigmatized nature of the subject matter) and also to the unreliability and distortion of retrospective recall on the part of the subjects. It is argued that these biases are liable to lead to systematic under-estimation of rates of relevant risk activity in the general population.
(McQueen et al., 1989; Turner et al., 1995) and to inaccurate and cognitively distorted estimates of sexual activity in general (Brewer, 1988; Linton, 1986).

The issue of reactivity can to a considerable extent be overcome by removing the data collection from the face-to-face interview and relying either on an interaction which is not visual (e.g. computer-assisted telephone interviewing (McQueen et al., 1989; Tieleman, 1990, or audi- or telephone-audio computer assisted self-interviewing (O’Reilly et al., 1994; Turner et al., 1997)) or adopting the simpler expedient of using the self-administered questionnaire (SAQ), which in both sexual behaviour studies and in cognate areas such as drug-taking, has been shown to be a less biased method (Sundet et al., 1990; Turner et al., 1992).

Cognitive biases also come into play in most survey methods, especially when subjects are required to count instances of their behaviour, primarily due to factors of memory and recall. How accurately can subjects recall events when called upon to do so some time after their occurrence and what biasing processes come into play when subjects make estimates of rates or frequencies of occurrence? Retrospection has been paid especial attention, both from the point of view of systematic processes of bias in autobiographical recall—forgetting, telescoping, chunking (Croyle & Loftus, 1993; Forsyth et al., 1992; Neisser & Winograd, 1988; Rubin, 1986), together with strategies of aggregation such as rounding and 'glossing up' (Coxon, 1988a). Others have shown how these biases show themselves in behavioural research (Coxon, 1988b; Janson, 1990; Weinhardt et al., 1998; Wu et al., 1988). Retrospection bias is probably the greatest threat to accuracy in sexual behaviour data. But it is not entirely clear how such bias may best be combated, and what other methods might be appropriate for reducing it.

The task of improving the accuracy of measures may suggest that there is some objective ‘gold standard’ which acts as criterion for assessment of their validity. Ideally, but not realistically, the appropriate tertium quid would appear to be the direct (and not laboratory-based) observation of the sexual behaviour, but even this is problematic and yet without this there can be no unquestioned criterion of validity. Failing this, the notions of relative precision of a measure and the concordance or convergence of several measures may serve as alternatives. The diary method has been cited as a suitable alternative measure for sexual behaviour in several reviews: Catania et al. (1990, p. 174) cite the greater precision of the diary for sexual activity in several reviews: Catania et al. (1990, p. 238) suggest that the concordance of estimates from SAQs and coital diaries would be sufficient to validate each method, and allow the strengths of each to be capitalized upon.

Diaries are used increasingly in social science generally (Butcher & Eldridge, 1990; Corti, 1993; Richardson, 1994; Verbrugge, 1980) and especially in sexual research (Coxon, 1996a; Coxon & Coxon, 1993b; Graham & Bancroft, 1997). Although diaries are subject to a number of shortcomings, especially potential selection bias and difficulties in use for large populations, the advantages of diaries are also very considerable (Catania et al., 1990; Coxon, 1988a). Most importantly, the short time period between the sexual act and its recording makes for greater precision and minimizes retrospective error (Bancroft, 1997; Berk et al., 1995; Coxon, 1992; 1994; Dex, 1991; Freeman et al., 1996; Hilton, 1989; James et al., 1991; Juster, 1985; Kalfs & Saris, 1990; McLaw & Lows, 1990; Poikolainen & Karkkainen, 1993; Reading, 1983; Robinson, 1985), and diaries also allow more detailed, contextually-specific and naturally expressed accounts of behaviour. It would appear, then, that if it were possible to show convergence between two self-administered methods, the SAQ and the self-administered Sexual Diary (SASD), assessment of sexual behaviour rates would be more firmly, accurately and validly demonstrated.

But the direct comparison of SAQs and SASDs is no simple matter, and the
few studies which have attempted to show convergence of diary and survey estimates (Fortenberry et al., 1997; Gold & Rosenthal, 1995; McLaws et al., 1990; Phellas, 1994; Poikolainen & Karkkainen, 1993) have tended to suffer from methodologically serious shortcomings of one sort or another, and in particular (i) that the data referred to two different reference time periods, making it possible that the behaviour was different in those intervals, or (ii) comparison was only made of aggregate rates rather than on differences at the individual level.

Whilst the relative agreement between questionnaires and diaries has been shown (Huysman, 1996; McLaws, 1990; Reading, 1983), these authors also report that estimates derived from diaries were generally lower than estimates reported in the interview or questionnaire (Huysman et al., 1996; McLaws et al., 1990; Phellas, 1994). In a study most similar to the present one, McLaws et al. (1990) report that discrepancies also differed systematically by the activity involved, not only in terms of the size of discrepancies, but in some cases also the direction (sign) of the discrepancies. Moreover, with specific reference to sexual behaviour, they report (ibid., pp. 275–276) that diary and interview estimates are in good agreement for more infrequent (homo)sexual activities, but in poor agreement for frequent ones, and an earlier SIGMA study (Phellas, 1994) found that the only significant estimate which was lower in the interview than in the subsequent month’s diary data was receptive anal intercourse—an ominous finding, consistent with McLaws et al., but suggesting that interviews (in this case) were least accurate when reporting rates of the most significant risk behaviour, and that such under-reporting was systematic.

In this paper three questions are therefore addressed:

- Is there agreement at the profile (aggregate) level between the two methods (SAQ and SASD) when estimates are made referring to the same time period?
- How sure do the respondents feel about making their retrospective questionnaire estimates of the frequency of the sexual behaviour which has been previously recorded in their diaries?
- At the individual level: (a) are there systematic tendencies to under-over-report? and (b) are divergences greater for some sexual acts than for others?

**Method**

The research reported here is part of a set of validation studies of the method of sexual diaries by Project SIGMA, Essex. The sexual diary method has been developed and used since 1984 in conjunction with interview and serological testing of gay and bisexual men in England and Wales (Coxon, 1988a; 1992; 1994; 1996a; 1996b).

The standard Project Diary Kit was used to obtain the basic SASD diary data (Coxon, 1992; SIGMA, 1993; 1996a). The diary consisted of four one-week pages, where the diarist wrote his entry on a daily basis for each sexual 'session', being reminded on each page about the characteristics of the sexual behaviour which were to be recorded. Sexual partners were entered as a numbered list, together with their characteristics, such as age and HIV status.

**Subjects**

Recruitment to the Project SIGMA Sexual Diary studies has been via the five-wave Longitudinal Panel (Coxon, 1996a; Davies et al., 1993), supplemented by biennial appeals for volunteers, made in the British gay press. The data for this study were obtained in the 1992/3 study. Following an appeal in three gay publications for volunteer diarists, 178 men wrote
in for a diary kit. Of these, just under one-half (86) completed and returned their month-long sexual diary. Immediately the completed diary was returned, they were sent (without prior warning) a new letter and form asking them to estimate the number of times they had done each of a set of 14 sexual acts during the period of their diary, together with a request to give a five-point rating of how certain they were of each estimated number. An incentive of £2 was offered, to be donated to the AIDS charity Terrence Higgins Trust on successful completion and return. Seventy-four of the 86 diarists (86\% participation rate) sent back a set of estimates of frequency of behaviour and associated certainty ratings. The purpose of this supplementary exercise was to obtain estimates of their sexual activity in self-administered questionnaire format from the same men who had completed (and returned) their sexual diary, and referring to the same time period.

Measures

The three main measures are:

- The diary count: obtained by counting post factum (by means of specially-written software (Coxon & Coxon, 1993c)) the number of times that each of the 14 sexual acts occurred in the subjects’ encoded diaries.
- The corresponding (retrospective) questionnaire estimate: made by each subject of the number of times in the diary-month they had done each of the acts, together with
- The associated five-point graded certainty rating: of how sure the diarist was about the numerical estimate of each behaviour, which he had just made in the questionnaire.

The three main measures refer to the same set of 14 sexual acts. The sexual acts consist of:

1. The active and passive variants of four homosexual behaviours: three common acts carried to ejaculation: masturbation, fellatio and (unprotected) anal intercourse together with one relatively uncommon act ("rimming"—anilingus or oral-anal contact).
2. Active and passive variants of fellatio where the inserter ‘pulled out’ before ejaculating.
3. Active and passive anal intercourse with the use of a condom.
4. Solitary masturbation to ejaculation, together with
5. Vaginal intercourse (which does not occur in these data).

These particular acts (specified in Figure 2, column vii) were chosen to vary acts of different frequency (common versus uncommon) (i) to include acts of questionable (and currently questioned) safety (ii), to include high-risk behaviours with and without protection (iii), to include the single most common activity (iv) and the most significant heterosexual activity (v). For purposes of simplification, this set is reduced to 11 in subsequent analysis and is specified in Figures 1 and 2.

Discrepancy measure

In order to examine the discrepancies between a subject’s diary and questionnaire estimates, a simple measure of difference was used (D–Q), defined by subtracting a subject’s retrospective questionnaire estimate (Q) of a given sexual act from the corresponding count derived from his diary record (D). If this value is zero, the diary count and the questionnaire estimate are the same; if negative, the questionnaire estimate is the higher and if positive, the diary count is the higher.
RESULTS

In the following analysis, the 11 sexual acts will be presented in 'risk order'—risk to the subject, that is. The risk order is induced by imposing three constraints:

1. **behaviour**: that anal intercourse > (is more risky than) fellatio > oral–anal > masturbation;
2. **modality**: that ejaculation of semen 'in' the diarist (receptive anal and oral intercourse) > ejaculation 'on' the diarist;
3. **prophylaxis**: no condom use > condom use.

The resulting full rank order$^{10}$ of risk is given in column (i) of Figure 2. Such a risk order is used to recognize how the main variables vary with risk.
<table>
<thead>
<tr>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
<th>(v)</th>
<th>(vi)</th>
<th>(vii)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk rank</td>
<td>Diary count (D)</td>
<td>Questionnaire estimate (Q)</td>
<td>Relative difference (D-Q)/D</td>
<td>Certainty rating</td>
<td>SIGMA code</td>
<td>Sexual act</td>
</tr>
<tr>
<td>11 high</td>
<td>0.65 (2.08)</td>
<td>1.40 (3.15)</td>
<td>-1.15</td>
<td>4.42 (1.28)</td>
<td>PF.M</td>
<td>Were you fucked by a guy?</td>
</tr>
<tr>
<td>10</td>
<td>0.54 (2.15)</td>
<td>1.08 (3.23)</td>
<td>-1.00</td>
<td>4.42 (1.27)</td>
<td>AFH.</td>
<td>Did you fuck a guy?</td>
</tr>
<tr>
<td>9</td>
<td>1.05 (2.15)</td>
<td>1.34 (3.06)</td>
<td>-0.28</td>
<td>4.11 (1.45)</td>
<td>PF.C</td>
<td>Were you fucked by a guy?</td>
</tr>
<tr>
<td>8</td>
<td>0.76 (1.75)</td>
<td>1.44 (3.65)</td>
<td>-0.89</td>
<td>4.13 (1.32)</td>
<td>AS.M</td>
<td>Did you suck a guy?</td>
</tr>
<tr>
<td>7</td>
<td>0.49 (1.35)</td>
<td>0.67 (1.59)</td>
<td>-0.37</td>
<td>4.04 (1.40)</td>
<td>PSH.</td>
<td>Were you sucked by a guy?</td>
</tr>
<tr>
<td>6</td>
<td>0.49 (1.37)</td>
<td>0.73 (1.70)</td>
<td>-0.52</td>
<td>4.34 (1.31)</td>
<td>AFC.</td>
<td>Did you fuck a guy?</td>
</tr>
<tr>
<td>5</td>
<td>2.75 (4.05)</td>
<td>3.75 (6.47)</td>
<td>-0.36</td>
<td>3.22 (1.40)</td>
<td>PWX.</td>
<td>Were you wanked by a guy?</td>
</tr>
<tr>
<td>4</td>
<td>2.41 (3.40)</td>
<td>3.75 (5.18)</td>
<td>-0.56</td>
<td>2.97 (1.38)</td>
<td>AWX.</td>
<td>Did you wank a guy?</td>
</tr>
<tr>
<td>3</td>
<td>0.63 (1.34)</td>
<td>1.01 (1.77)</td>
<td>-0.60</td>
<td>4.14 (1.35)</td>
<td>ARG</td>
<td>Did you rim a guy?</td>
</tr>
<tr>
<td>2</td>
<td>0.49 (1.23)</td>
<td>1.09 (1.90)</td>
<td>-1.22</td>
<td>4.05 (1.35)</td>
<td>PRI.</td>
<td>Were you rimmed by a guy?</td>
</tr>
<tr>
<td>1 low</td>
<td>10.93 (9.64)</td>
<td>14.37 (13.07)</td>
<td>-0.31</td>
<td>3.07 (0.85)</td>
<td>SWX.</td>
<td>Did you wank yourself?</td>
</tr>
</tbody>
</table>

Fig. 2. Summary diary count, retrospective estimate and certainty ratings data by risk order.

Do the diary and questionnaire estimates differ?

The relevant summary data relating the diary counts and the questionnaire estimates are illustrated in Figure 1 and the summary data are given in Figure 2, arrayed in order of risk. The most striking characteristic of the profiles of SAQ and SASD average frequencies is that, as in previous studies, diary counts are lower than questionnaire estimates, and are so for all sexual acts. The largest average differences are not surprisingly for the most common activities (masturbation), and the biggest relative differences (relative to the frequency of occurrence, see column iv) include the most risky activity: unprotected anal intercourse. But the overall profiles are very similar indeed ($r = 0.99$), though this masks some rank differences ($r = 0.76$).

How sure are subjects about their estimates?

Several suggestions have been made about the likely accuracy of recall of sexual data which may be illuminated by the certainty ratings data. The most usually cited hypothesis is:
• That less common acts are recalled more accurately than more common ones (due presumably to the greater ‘surprise value’ of rare acts).

The relative frequency with which homosexual acts occur in sexually active gay populations is basically the same, whether the reference is to the number of acts or the number of individuals who do it (Coxon, 1969a, pp. 64–86; Davies et al., 1993, p. 106), and this makes it simple to define the ‘more common acts’ as masturbation and oral sex. This hypothesis holds quite well: the less common the act, the more sure subjects are of its accuracy ($r = -0.68$, $\tau = -0.43$, between averages).

More relevant from the point of view of making epidemiological predictions is the question of whether certainty of estimates is related to the riskiness of the sexual act: is it the case that the more risky the act, the higher the certainty of the estimate? Again the answer is yes ($r = 0.15$) and although the correlation of average certainty and average riskiness is only modest, the Multisample Median test and Jonckheere’s Trend test give highly significant ($p < 0.001$) differences in certainty ratings between the sex act medians. Thus diarists are most sure about their estimates of those sex acts which are of the highest risk. In particular (following the median judgements),

• They are virtually all certain how often they had engaged in anal intercourse (both insertive and receptive anal intercourse, whether or not protected).
• They are relatively unsure about how often they had engaged in fellatio.
• They are only moderately sure of the number of times they engaged in masturbation with others, and not at all sure of their estimate of solitary masturbation.

Analysis of individual differences

Having confirmed that the profiles of the average frequencies of sexual acts for the diary and questionnaire methods are relatively similar, it is important now to examine discrepancies between questionnaire estimates and diary counts at the individual level.

First, what is the gross overall pattern of discrepancy between diary count and questionnaire estimate (summed over all individuals and acts)? It is negative and averages just over one act, i.e. there is an overall tendency for individuals to give questionnaire estimates which are higher than diary counts, as there is in the aggregate data.

To what extent do individuals differ in their discrepancies? Calculating the average discrepancy score for each individual (over all the sex acts he engages in), the distribution of individual differences (presented in Figure 3) is clearly skew, dominated by rather small differences around zero (median: $-0.75$, IQR: 2.2) and there is a long tail toward negative values. This takes no account of the fact that the individual sexual repertoires differ considerably in the range of acts they include, nor of the differences between the acts: the overall set is highly dominated by masturbation.

In order to gain further insight about how discrepancies are distributed, it is useful to define the typology of discrepancy scores (D–Q), as shown in Figure 4. Irrelevant discrepancies are simply included for completeness of the typology. ‘Mistaken’ and ‘forgotten’ discrepancies, where one or other frequency is zero, are presently ignored in calculations. Applying the typology first to the sum total of discrepancies, 55% are over-reports, 20% are identical and 25% are under-reports, i.e. twice as many judgements are over-estimates than are under-estimates, and only one-fifth of questionnaire estimates are identical to diary counts.

The usual individual tendency in discrepancies is mostly in one direction—predominantly
FIG. 3. Overall individual differences between diary and questionnaire estimates (over all sexual acts).

To under-report or predominantly to over-report. Very few indeed have a mixture of under- and over-estimation—almost all err in one direction only, and by far the majority of discrepancies are negative. This shows quite clearly that:

- There is a strong tendency for an individual to be either an under-estimator (43%) or an over-estimator (26%).
- That few (17%) have a mixture of positive and negative discrepancies.
- That very few individuals indeed give consistently identical estimates—only one person has zero discrepancies for every sexual act he engages in—and few (14%) are even close to zero, and they have a very restricted sexual repertoire.

<table>
<thead>
<tr>
<th>Pattern of discrepancy</th>
<th>Description of discrepancy</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>[D,Q]</td>
<td>(The sexual act concerned ...)</td>
<td>Irrelevant</td>
</tr>
<tr>
<td>[0,0]</td>
<td>Does not feature in diary and zero estimate in questionnaire</td>
<td></td>
</tr>
<tr>
<td>[0,*]</td>
<td>Does not feature in diary, but questionnaire estimate is over zero</td>
<td>Mistaken</td>
</tr>
<tr>
<td>[*,0]</td>
<td>Diary count over zero, but questionnaire estimate is zero</td>
<td>Forgotten</td>
</tr>
<tr>
<td>[D &gt; 1]</td>
<td>Diary count is greater than questionnaire estimate</td>
<td>Under-report</td>
</tr>
<tr>
<td>[D = Q]</td>
<td>Diary count is same as questionnaire estimate</td>
<td>Identical</td>
</tr>
<tr>
<td>[D &lt; Q]</td>
<td>Diary count is less than questionnaire estimate</td>
<td>Over-report</td>
</tr>
</tbody>
</table>

(*) signifies a value over zero.

FIG. 4. The typology of discrepancy scores.
**Table 1. Individuals' sexual behaviours: types of discrepancies between diaries and questionnaires**

<table>
<thead>
<tr>
<th></th>
<th>Did not do in month (%)</th>
<th>Mistaken/ forgotten</th>
<th>Under-report</th>
<th>Equal estimates</th>
<th>Over-report</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masturbation (with others) (p &lt; 0.01, paired t-test)</td>
<td>24</td>
<td>31</td>
<td>40</td>
<td>14</td>
<td>15</td>
<td>100</td>
</tr>
<tr>
<td>Fellatio (NS)</td>
<td>65</td>
<td>30</td>
<td>34</td>
<td>22</td>
<td>14</td>
<td>100</td>
</tr>
<tr>
<td>Anal intercourse:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with condom (NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>active, no condom</td>
<td>68</td>
<td>30</td>
<td>12</td>
<td>36</td>
<td>22</td>
<td>100</td>
</tr>
<tr>
<td>(p &lt; 0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>passive, no condom</td>
<td>80</td>
<td>36</td>
<td>50</td>
<td>14</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>(p &lt; 0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>31</td>
<td>11</td>
<td>26</td>
<td>32</td>
<td>100</td>
</tr>
</tbody>
</table>

Finally, it remains to see the ways in which the pattern of an individual's discrepancies are related to particular sexual acts. The profile of discrepancies for a number of relevant sexual acts is given in Table 1. The patterns for each act are quite distinct:

- **Masturbation**: is typified chiefly by under-reporting; little difference between active and passive modalities; diary counts are significantly higher than questionnaire estimates (p < 0.001).
- **Fellatio**: Somewhat more under- than over-reporting; slightly more difference between modalities. Diary counts usually greater than questionnaire estimates (NS).
- **Anal intercourse**:
  - with condom: Highest identical reporting; little difference between modalities. (NS differences between diary counts and questionnaire estimates.)
  - no condom: Very different pattern compared to condom-protected-anal intercourse, and major differences between active and passive modalities:
  - active: Very large degree of under-reporting; no over-reporting.
  - passive: Considerable identical and over-reporting.

Masturbation and fellatio thus present a distinct pattern of discrepancies: questionnaire estimates are usually lower than diary counts. But the major, significant and important discrepancies in reporting occur for anal intercourse, especially for its riskiest unprotected variants, and for both its active and passive modalities. Compared to these, the discrepancies for other behaviours are unimportant. The pattern which is most distinctive and different is that for anal intercourse without a condom, which is highly dissimilar both to anal intercourse with a condom and to receptive anal intercourse without a condom. Here is the Achilles' heel of the SAQ compared to the SASD.

**Conclusions and discussion**

The diary method has been used as an appropriate method for collecting sequential and time-linked data on sexual behaviour since 1982, and has proved to be a reliable tool for the study of sexual behaviour in a number of pioneering studies, resistant to interview and retrospection bias and yielding detailed and often unique information. But it is undoubtedly a demanding method (of both respondents and analysts), subject to selection bias and (as it stands) it is unlikely to be a substitute for survey methods.
The study reported here compares two methods for measuring the sexual behaviour of gay men—self-administered diary and subsequent questionnaire estimates—both of which refer to the same index month. The analysis confirms that the sexual diary method is likely to be more accurate than retrospective questionnaire data, but that they are strongly related. In particular:

1. Diarists attribute highest certainty in their questionnaire estimates of sexual behaviour to the less common sexual acts, and in particular for higher-risk acts which involve anal intercourse. This is confirmed in the smaller relative discrepancies associated with these same acts.

2. At the individual level, post factum questionnaire estimates also yield much the same aggregate profile as diary counts, but they consistently yield higher average frequencies.

3. Most individuals have a consistent pattern of over-identical/under-reporting estimates across different sexual acts.

4. The main important exception occurs for the highest-risk sexual activity: anal intercourse without a condom which differs markedly according to whether it refers to insertive (active) or receptive (passive) variants. The discrepancies may in part reflect a sort of ‘social desirability’, or gendered, effect: active partners tend to under-estimate, and passive partners tend to over-estimate the amount of highest-risk sex in their questionnaire accounts. If the diary is taken as the more precise method, SAQ estimates of insertive anal intercourse would exhibit a considerably lower incidence and estimates of receptive anal intercourse would assume a higher incidence than is the case.

How representative or generalizable are these results and how do these results apply to the issue of obtaining better data on sexual behaviour? First, issues of representativeness. The answer depends in part on known facts about samples of gay men:

- That prevalence is hard to establish in random sample studies due to interviewer effect, covert practice and adjudged deviance of homosexuality (Johnson et al., 1994, pp. 189–193; Kalton, 1993).

- That samples of gay men based upon volunteer subjects are known to be subject to over-representation of younger men, of those living in the metropolis, of those overt about their sexuality and of those with higher educational qualifications (Coxon, 1995; Davies, 1986; Fitzpatrick et al., 1990; Hunt et al., 1991).

- That within the SIGMA samples, those completing SASDs are likely to be slightly more sexually active (an empty diary or one consisting simply of solitary masturbation is unlikely to be volunteered or returned), and not be in a regular relationship (Coxon, 1990; 1996a).

The current sample is volunteer based and is small in size, but it is at least as representative as the main SIGMA sample and earlier diary studies, and differs primarily in being smaller in size, almost certainly due to ‘volunteer fatigue’, since the numbers of those responding to Project SIGMA appeals has declined sharply from 1987 to the present. However, when the aggregate sexual behaviour profiles from the various sexual diary data sets are compared, whether emanating from sub-samples within the SIGMA panel (Coxon, 1996a; Phellas, 1994) or from the biennial appeals—they yield markedly similar results (compare, for instance, Coxon, 1994, pp. 136–137). More relevantly, the finding that diary data estimates of rates of the main sexual behaviours are in aggregate systematically lower than questionnaire data holds for all SIGMA data sets where it is possible to compare interview
data with diary data for the same panel members, even though the reference month may be different. The earlier study of divergences in SIGMA panel members' diary and interview data (Phellas, 1990) had also detected that estimates of anal intercourse exhibited systematic individual differences in reporting, relative to other behaviours, but the fact that estimates for receptive and insertive variants of unprotected anal intercourse were biased in opposite directions was not discovered until the present analysis.

In the preliminary section, reasons were adduced as to why SASDs might be considered to give more accurate estimates than SAQs, but it was stated that, in the absence of a behavioural *tertium quid* for homosexual behaviour, no determination of greater accuracy is possible. But the results presented here can also be used the other way round—in effect, SASDs can be used to validate or check on survey data. It would be difficult, costly and probably impractical to use the method of SASDs to make national or large-scale studies of sexual behaviour, and the SAQs and interviews are much better adapted to that. However, if a random sub-sample of such a large-scale study were to use SASDs, it would be possible not only to check on patterns of under-over-reporting, but also to examine the contexts in which sexual behaviour (and especially risk behaviour) take place, and which have specific relevance to explaining it. And this is, after all, a major advantage of the method of SASDs.

Notes

[1] Recording real-time sexual activity has major problems of reactivity (if observers are present), and of selection bias if recruited from volunteers, and the behaviour is in any event continuous in form, rather than being re-interpreted into categorical form, as self-reports of sexual behaviour invariably are.

[2] The data were a balanced random sample from the core SIGMA data set, consisting of 1,035 month/diaries obtained from panel members drawn from ten sites throughout England and Wales in the period 1987–1993.

[3] Project SIGMA (Socio-sexual Investigations of Gay Men and Aids). The research in this paper was funded by the Medical Research Council and by the Department of Health (UK). Work on Sexual Diary software and on the validity and reliability of the sexual diary method reported here is funded under a series of grants from the Department of Health and reported in Coxson (1996a; 1996b). The views expressed in this paper are those of the author and not necessarily those of the Council or of the Department.

[4] These read: ‘Remember, each session should include: The Time, The Place, The Partners (from partner list). Then, describe the session in your own words. Remember to mention exactly what happened to the “come” (ejaculate) and always mention the use of condoms. List any accompaniments you use (poppers, lubricants, drugs, sex toys ... ) and mention how much you drank each day if it is associated with sex.

[5] Proportionate responses were: from BOYZ (64%), Capital Gay (18%) (both weekly free newspapers, the latter now defunct) and Gay Times (17%) (monthly). The first has a younger readership and is primarily London-based and ‘scene’-oriented, the second is also London-based, but more serious and with a slightly older readership, and the third has a more substantial and long-established national readership.

[6] The wording of instructions and format of the questions were as follows:

HOW OFTEN do you reckon you did each of the following sexual acts?

( write in the box the number of times you estimate )

... then for each number, underline the category below that best indicates how sure you are.

In the month of the diary, how often ...

1. Wanking a guy till he comes

(I am: CERTAIN/VERY SURE/MODERATELY SURE/NOT VERY SURE/NOT AT ALL SURE of this number)

[7] The term ‘modality’ is used to differentiate active and passive variants of the same behaviour. ‘Active’ refers to ego doing an act to *alter*, ‘passive’ to *alter* doing an act to ego; it is not identical to the insertor/insertee distinction.

[8] Masturbation, fellatio and anal intercourse together account for an estimated 81% of homosexual acts (Coxson, 1996a, p. 167), anilingus accounts for 2%.

[9] The acts of fellation referring to ‘pulling out’ were found to be ambiguous with respect to ejaculation and were removed from further analysis. (For instance: ‘Sucking a guy: he pulls out before he comes’—does *alter* ejaculate}
after withdrawing, or does he withdraw in case he ejaculates?) Vaginal intercourse was not reported and was also removed.

[10] The order cannot be reduced to a strict order, due to the relatively arbitrary placing of oral-anal contact (ARI and PRS), where there is considerable risk of transmission of Hepatitis B, but little of HIV.

[11] Using the Index of Dissimilarity (Blaauw & Duncan, 1967), the most similar profiles are for masturbation and fellatio (8%) and for receptive anal intercourse without a condom and anal intercourse with a condom (16%).

The maximum dissimilarities are for insertive anal intercourse without a condom and anal intercourse with a condom (44%) and (insertive anal intercourse without a condom and receptive anal intercourse without a condom) at 44%.

[12] Regression of diary counts on interview estimates (not presented here) also shows that active/insertive and passive/receptive variants work in different ways; insertive anal intercourse discrepancies are basically higher by a constant, whereas receptive anal intercourse (the more risky behaviour) discrepancies differ by a multiplier.

[13] I am grateful to an anonymous referee for making this point.

[14] Altruism is a potent incentive to participate by those who are already motivated to do so; monetary incentives are more successfully used to persuade those who would not otherwise participate (see Church, 1993).

References


How Many Account for How Much? Concentration of High-Risk Sexual Behaviour Among Gay Men

Anthony P. M. Coxon
University of Essex, England

Thomas J. McManus
The Royal London Hospital, England

The data set of 2,182 gay/bisexual men's month-long sexual diaries collected by the United Kingdom Socio-sexual Investigations of Gay Men and Aids (Project SIGMA) was used to analyse the extent to which acts of anal intercourse (AI) are distributed among gay men, using the Lorenz/Gini methodology for measuring concentration. Most individuals (60%) who engage in AI do so only once or twice a month, but there is a long tail of those who do it much more. In terms of the amount of AI acts, one tenth of the individuals are performing half of the acts of AI. The Gini coefficient of concentration is high (0.55). Factors most affecting rates and concentration of risk behaviour include relationship status, HIV sero-status and concordant/discordant partner status.

The number of individuals who engage in risk-prone sexual behaviour is of obvious and primary importance in the analysis of Aids/HIV infection. Anderson (1988) points out that the degree to which this infection will spread is basically dependent on the magnitude of the basic reproductive rate, but that a small proportion of individuals with high rates of partner change (and hence of risk-prone sexual behaviour) make a disproportionate contribution to the reproductive potential of the infection. Hence, the amount of such risk activity is equally important as the number engaging in such activity and, taken together, they define the concentration of risk-prone sexual behaviour.

One reason why concentration of sexual behaviour has received relatively little attention in the literature is due to the difficulty of obtaining reliable and accurate measures of the actual amount of risk-prone sexual behaviour (at least by the use of conventional methods of data collection). As an antidote to recall bias in surveys (Baddeley, 1979; Coxon 1993a, 1993b; Dex, 1991), self-completion methods have been proposed which record the events simultaneously, or only a short period after they happened. In particular, diary methods have been used successfully for studying sexual behaviour in general (Fortenberry, Cecil, Zinets, & Orr 1997; Freeman, DeRubeis, & Rickels 1996; Graham & Bancroft 1997; James, Bignell, & Gillies 1991; Leigh, Gillmore, & Morrison 1998; Reading, 1983) and in studies of gay men in particular (Coxon, 1994, 1996; Coxon & Coxon, 1993a, 1993b; Gold & Rosenthal, 1995; McLaws et al., 1990). Moreover, diaries have been found to produce data which are as reliable as, and in most cases more valid than, questionnaire procedures (Conrat, Higgins, & McLean 1993; Coxon, 1996, 1999; Freeman et al., 1996; James et al., 1991; Juster, 1985; Leigh et al., 1998; McLaws et al., 1990; Phellas, 1994; Poikolainen & Karkkainen, 1993; Robinson, 1985; Weinhardt, Forsyth, Carey, Jaworski, & Durant, 1998).

In this paper we rely upon the data derived from month-long sexual diaries collected by Project SIGMA in England and Wales (Coxon 1996) in order to study various types of homosexual anal intercourse (AI) and to do three things:

1. To integrate information on incidence (numbers of men engaging in AI) and rates (amount of AI being done) by means of the Lorenz Diagram, and to examine the concentration of such risk behaviour using the associated Gini coefficient.

2. To see if different variants of high-risk behaviour (e.g., protected vs. unprotected anal intercourse) differ in concentration.

3. To see what variables affect concentration.

METHOD

Research Instrument: The Sexual Diary

The method of sexual diaries (MSD), was developed and used within Project SIGMA from 1984.1 Diaries are written in natural English (but according to a structured schema),

1The sexual diary method has been developed in conjunction with interview and serological testing of gay and bisexual men in England and Wales, by Project SIGMA (Coxon, 1988, 1994, 1996, 1999).
and are filled in on a daily basis. The method yields highly contextualised and detailed information of the sequence of sexual activities (Parker & Carballo 1990). The diary format is based on the unit of the sexual session (consisting of a sequence of sexual acts with a natural time-marked start and end), in turn consisting of sexual acts, each of which includes specific information on agency (what we call modality—what does who to whom), on the sexual behaviour itself (e.g. anal intercourse), and on the precise destination of the ejaculation of semen (if it happens) with respect to both partners.

Data Set

The sexual diary data-set used in this paper consisted of the 2,182 individual diary/months of 1,035 men who have sex with men, drawn from ten locations/sites in England and Wales in the four waves of Project SIGMA between 1987 and 1992. Most of the analysis focused on the subset of 628 diary/months which included one or more acts of anal intercourse in this period.

Measures

The three variables of interest are:

1. The amount of risk-activity, measured by the number of sessions which contain one or more risk-acts (identified and extracted from the diary record);

2. The number of men engaging in a given risk-activity such as anal intercourse. Information is obtained by counting the number of diarists who engaged in the activity during the standard month period, and is also extracted from the diary-record itself; and

3. The riskiness of a sexual session, measured in terms of what combination of risk-factors occurred. The risk factors measured were whether anal intercourse (AI) occurs in a session, whether ejaculation (E) occurs as a result of the act of AI, and whether a condom is worn in the act of anal intercourse.

A typology of increasing risk-activity is presented in Table 1. This simple typology differentiates risk sessions (i.e., those involving AI) in terms of whether ejaculation occurs and, if so, whether or not it is protected. In brief, high-risk sessions are any which involve ejaculation, and highest-risk sessions are differentiated in terms of whether or not the ejaculate is into a condom (C: protected) or into the receptive partner himself (D: unprotected).

Measuring Concentration

The Lorenz curve and the associated Gini coefficient is a method developed primarily within Economics (see Cowell 1995; Stiglitz, 1993) to measure inequality and the concentration of ownership in the distribution of resources in general, and of income, wealth, and property in particular. Subsequently, it has been extended to deal with inequality and concentration in other domains, such as political power and resources (Alker, 1965), geographic and regional concentration (Shelburne & Bednarzik, 1993; Smith, 1979), and health (Davey Smith & Egger, 1996). In this paper, the concentration of sexual risk activity is measured using this same methodology.

The Lorenz Diagram gives a visual representation of inequality and concentration in a single graph. Its use in the context of sexual behaviour can be illustrated by reference to the following (fictitious) set of data (see Figure 1). Suppose we have found the following empirical information about the distribution of sexual acts (having sorted the

<table>
<thead>
<tr>
<th>Table 1. Typology of Risk Sessions</th>
</tr>
</thead>
</table>
| Anal
| Intercourse
| Ejaculation
| Condom
|------------------------------------|
| A: Risk session ✓
| B: High risk session ✓
| C: Highest risk session (condom-protected) ✓
| D: Highest risk session (unprotected) ✓

Legend:

✓ must contain
* may or may not contain
X must not contain

Figure 1. Example of a Lorenz Curve (fictitious data).

The analysis is done using the LORENZ program in the SDA suite of programs (Coxon & Coxon, 1993b).
data by those who do exactly 1 sex act, exactly 2 acts, and so on; the first 50% of the men do 18% of the sex acts, the first 75% of the men do 30% of the sex acts, the first 90% of the men do 50% of the sex acts, and 100% of the men do 100% of the sex acts. This provides the basic information which the Lorenz diagram portrays (see Figure 2). The diagram is designed to display what proportion of the population "owns" what proportion of the "scarce resource" by charting what proportion of the population (horizontal axis) has engaged in what proportion of the total number of sexual acts (vertical axis). The data are plotted for 1 act, up to 2 acts, up to 3 acts, and so on, and the Lorenz Curve (guaranteed by its cumulative nature to be monotonic upward), formed by joining the points, is portrayed in Figure 1.

This Lorenz curve is then compared to what would be the case if sexual acts were shared equally, so that 10% of the population did 10% of the sexual acts, 20% did 20% of the sexual acts, and so on. This forms the line of equality (or equal distribution) moving up diagonally at 45° from the bottom left to the top right of the chart, portraying the baseline "equality" distribution on what would be the case if the amount of sexual activity were the same for all. The difference between these two curves gives a visual representation of the relative inequality, or of concentration—how far from equality the distribution it is, and how far the sexual activity is concentrated in a small proportion of individuals. A numerical measure of overall inequality or concentration is given by the Gini coefficient, which measures the gap between equal distribution and actual distribution. The value of the Gini coefficient is 1 where there is total concentration/inequality (the very topmost fraction owns everything) and 0 if there is total equality or equal dispersion (everybody owns the same amount).

The Lorenz curve and associated Gini coefficient thus can be used in this case to measure how the stock of (risky) sexual activity is divided between the gay men involved, and can provide a useful numerical measure of the concentration of sexual risk. The substantive questions that can be addressed by this means are as follows:

1. What is the distribution over the specified levels of risk activity?

2. Are individuals equally liable to engage in unprotected variants? Or is high-risk sex concentrated in relatively few individuals? If so,

3. If high-risk sex is concentrated in relatively few individuals, what characteristics /contexts typify them?

---

\[ G = 1 - 2 \int L(u)du \]

and for a discrete variable with \( p \) categories sorted in order from the lowest to the highest, the coefficient is defined as:

\[ G = \sum \frac{X_{i}}{Y} \Delta Y, \]

where \( X_{i} \) is the cumulative percentage of \( X \) by unit, and \( \Delta Y = Y_{i} - Y_{i-1} \).

See also Alker (1965, pp. 41-42).

---

**Figure 2. Outcomes of acts of anal intercourse (from sexual diary data).**

<table>
<thead>
<tr>
<th>Potential Risk</th>
<th>Potential High Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>30% did not result in ejaculation of semen ([N]), and</td>
<td>Of these ejaculations,</td>
</tr>
<tr>
<td>70% involve ejaculation of semen. ((HR \text{ Protected}))</td>
<td>17% of ejaculations are into a condom ([C]) ((HR \text{ Protected}))</td>
</tr>
<tr>
<td>but 53% are unprotected. ((HR \text{ Unprotected})).</td>
<td>but 53% are unprotected. ((HR \text{ Unprotected})).</td>
</tr>
<tr>
<td>Of these unprotected ejaculations,</td>
<td>Of these unprotected ejaculations,</td>
</tr>
<tr>
<td>12% go &quot;elsewhere&quot; ([E])</td>
<td>12% go &quot;elsewhere&quot; ([E])</td>
</tr>
<tr>
<td>4% go on the partner’s body ([O/F])</td>
<td>4% go on the partner’s body ([O/F])</td>
</tr>
<tr>
<td>37% go into the partner’s anus ([H/M]) ((HR \text{ Highest Risk}))</td>
<td>37% go into the partner’s anus ([H/M]) ((HR \text{ Highest Risk}))</td>
</tr>
</tbody>
</table>

---

**RESULTS**

Analysis is restricted to data from those men who engaged in anal intercourse during the month of their sexual diary (one third of the total). ⑥

**Amount of Anal Intercourse**

First we examined the information about the amount, volume, or "stock" of the 2,481 acts of AI, without reference to the number of individuals involved. The outcomes are reported in Figure 2 as a Risk Hierarchy. This shows that if an act of anal intercourse occurs, it is most likely (70% of acts) to result in ejaculation; if ejaculation occurs, it is more likely to be without a condom (53% of all AI acts) than with a condom; if ejaculation occurs, it is most likely to be directly into the partner’s anus (37% of all AI acts), than elsewhere; and if ejaculation is into the partner’s anus, it is more than twice as likely to be without a condom (37% of all AI acts) than with a condom (17% of all AI acts). However viewed, these figures indicate that in this sample, anal intercourse to ejaculation without a condom into the partner’s anus (AI/NC) is the rule rather than the exception in gay men’s sexual intercourse.

**Concentration of Risk**

AI Baseline. Given this extent of highest risk behaviour, the question arises of how risk behaviour is concentrated, and whether the concentration increases as acts become more high-risk. To do this, three levels of the typology of Table 1 are considered: B: High Risk Session—anal

---

⑥In this 5-wave diary data set, two thirds of these (66%) contained no act of anal intercourse during the month concerned. The subset of interest is now the 628 (diary/months) which included one or more acts of anal intercourse.
Table 2. Distributions of Acts of Anal Intercourse (AI)

<table>
<thead>
<tr>
<th>N of acts of AI (per month)</th>
<th>N(diary/m)</th>
<th>%</th>
<th>Cum. %</th>
<th>Acts of AI</th>
<th>%</th>
<th>Cum. %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>244</td>
<td>38.9</td>
<td>38.9</td>
<td>244</td>
<td>9.8%</td>
<td>9.8</td>
</tr>
<tr>
<td>2</td>
<td>136</td>
<td>21.7</td>
<td>60.5</td>
<td>272</td>
<td>11%</td>
<td>20.8</td>
</tr>
<tr>
<td>3</td>
<td>62</td>
<td>9.9</td>
<td>70.4</td>
<td>186</td>
<td>7.5%</td>
<td>28.3</td>
</tr>
<tr>
<td>4</td>
<td>49</td>
<td>7.8</td>
<td>78.2</td>
<td>196</td>
<td>7.9%</td>
<td>36.2</td>
</tr>
<tr>
<td>5</td>
<td>34</td>
<td>5.4</td>
<td>83.6</td>
<td>170</td>
<td>6.9%</td>
<td>43.1</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>3.2</td>
<td>86.8</td>
<td>120</td>
<td>4.8%</td>
<td>47.9</td>
</tr>
<tr>
<td>7</td>
<td>11</td>
<td>1.8</td>
<td>88.5</td>
<td>77</td>
<td>3.1%</td>
<td>51</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>2.1</td>
<td>90.6</td>
<td>104</td>
<td>4.2%</td>
<td>55.2</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>1.1</td>
<td>91.7</td>
<td>63</td>
<td>2.5%</td>
<td>57.7</td>
</tr>
<tr>
<td>10–19</td>
<td>36</td>
<td>5.7</td>
<td>97.5</td>
<td>466</td>
<td>18.8%</td>
<td>76.5</td>
</tr>
<tr>
<td>20–49</td>
<td>14</td>
<td>2.2</td>
<td>99.7</td>
<td>405</td>
<td>16.3%</td>
<td>92.8</td>
</tr>
<tr>
<td>50–104</td>
<td>2</td>
<td>0.3</td>
<td>100</td>
<td>178</td>
<td>7.2%</td>
<td>100</td>
</tr>
<tr>
<td>Sums</td>
<td>628</td>
<td></td>
<td></td>
<td>2481</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Mean = 3.8; median = 2 acts of anal intercourse/month. IQR = 3.

intercourse (AI); C: Highest Risk Session (Protected)—ejaculation into a condom (AIE/C); and D: Highest Risk Session Unprotected—ejaculation into anus (AIE/NC). In so doing, attention can be directed to two main questions: Is high-risk sex concentrated in a few individuals or is it evenly spread, and are there characteristics which distinguish, say, the occasional practitioner from those for whom anal intercourse is the central part of their sexual repertoire? The basic information is given in Table 2.

Thus, 39% do only one act of AI per month, and this accounts for 10% of the stock of all 2,481 acts of anal intercourse. The distribution of men who do exactly $p$ acts of anal intercourse in a given month (Table 2, col. 2) is highly skewed toward the low end, with a median of 2 acts/month (mean of 4), and a long tail, which peters out to one man who engages in 104 acts in a month. The distribution of the number of acts (Table 2, col. 6) is quite different, and shows clearly that the bulk of acts of anal intercourse is concentrated in the higher reaches: A few men are accounting for a high proportion of the acts. These two constituent percentage distributions are presented (in the same scale) in Figure 3.

Figure 3a. Percentage of individuals who do exactly $x$ acts of AI a month. Results from SIGMA diary data.

Figure 3b. Percentage of all acts of AI accounted for by those who do exactly $x$ acts a month. Results from SIGMA diary data.
In Figure 3a, the first bar gives the percentage of the individuals who did exactly one act of AI a month, the second gives the percentage of those who did exactly two acts a month, and so on. The distribution is massively skewed and is dominated by those who engaged in AI only a few times a month, with a long tail consisting of the very few who engaged in a large amount of AI. Figure 3b presents the percentage of the acts that were accounted for by those who only engaged in AI once a month, twice a month, and so on, up to the amount accounted for by the individuals who engaged in AI a very large number of times. This distribution is also skewed, but is far less dominated by those who engaged in AI relatively rarely, and indicates how the few "extremists" account for considerable fractions of the acts of AI.

These two distributions are now put together in the Lorenz diagram in Figure 4. As the number of AIs/month goes up, the cumulative percentage of men (Table 2, col. 4) increases more rapidly than the increase in the cumulative percentage of AIs which this represents (Table 2, col. 7). So, when 10 AIs/month is reached, a full 93% of men but only 60% of the AIs have been accounted for. From 9 AIs up, the percentage of men unaccounted for dwindles, and beyond 20 AIs/month only 16 (2.3%) are left—but between them, this 2.3% of men are responsible for 23% of the AIs. Hence, the few in the higher reaches of the tail of the distribution contributed very disproportionately to the amount of anal intercourse: The top 10% of men account for almost half of the acts.7

The Gini coefficient for these data is 0.55, which is high on any account, comparing values of 0.20 to 0.45 for income inequality in Western democracies and 0.45 and 0.60 for developing countries (Deininger & Squire, 1996). A coefficient value of this size indicates that there is a high degree of concentration of potentially high-risk behaviour in a small fraction of gay men (or, equivalently, that anal intercourse is distributed in a highly unequal manner); moreover, anal intercourse is more highly concentrated than other sexual activity among gay men.8

This suggests that, with respect to concentration of anal intercourse, there are three groups of gay/bisexual men: (a) those who do not engage in anal intercourse, (b) those who do engage in a small amount of anal intercourse, and (c) a small group accounting for a high fraction of the activity.

How true is this? We know already that there is a sizeable group (about one third) who does not engage in AI, and there is also a sizeable group who does engage in AI, but usually between one and three times a month—less than once a week. For the first two groups, then, the description is basically accurate. The third group is more diffuse: There are certainly some men who engage in AI very frequently and who account for a significant fraction of the potential risk acts, but any exact point of division is bound to be arbitrary. If we take the top tenth, they may account for half of the acts of anal intercourse, but they range from the 2% who engage in AI 8 times a month, to the 0.15% who engage in AI 40 times a month and the one man who engages in AI 104 times a month. Those at the extreme end of the distribution cannot be dismissed as an amorphous, dispersed remnant. The sexual activity of even a few such individuals can have major epidemiological consequences for the transmission of HIV if they are HIV-antibody positive, if their anal intercourse is unprotected, or if they are primarily insertive partners. Are they "risk-rich"? One important clue9 is that those in the highest quartile are considerably more likely to carry a condom with them at all times (perhaps because they expect to engage in AI) compared to the lower quartile, who do not, and for whom the relatively rare AI "catches them by surprise," unprepared and unprotected.

Variants of Risk Acts

Does concentration increase for the variants of AI which are more directly implicated in HIV transmission? To answer this it was necessary to differentiate unprotected from protected anal intercourse, and to allow for differences arising from the modality (insertive/receptive role) of the act. Taken together, this defines four types of AIE, contrasting protected versus unprotected and insertive (active) versus receptive (passive) variants.10

In order to examine significant differences between these four types of AIE, it is appropriate to look at both the

---

7The full data show that the top 10%, 5%, 2%, and 1% of men account for [46%, 33%, 20%, 15%] respectively of the acts of AI.
8G = 0.48 for all sexual acts (Coxon, 1996, pp. 191-192).
9This information is obtained from the SIGMA interview data.
10In the sexual diary these are defined as [AF, H*], [PF, *M], [AF, C*], and [PF, *C].
Table 3. Concentration Characteristics of Four Types of AIE

<table>
<thead>
<tr>
<th>Summary characteristics</th>
<th>Type of AIE</th>
<th>With condom</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Active</td>
<td>Passive</td>
</tr>
<tr>
<td>No. of fuck-acts/month</td>
<td>434</td>
<td>442</td>
</tr>
<tr>
<td>No. of individuals</td>
<td>188</td>
<td>154</td>
</tr>
<tr>
<td>Average no. fuck-acts/month</td>
<td>2.31</td>
<td>2.87</td>
</tr>
<tr>
<td>Md = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dispersion (IQR)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>% who do exactly 1 fuck-act/month</td>
<td>54%</td>
<td>51%</td>
</tr>
<tr>
<td>% who do 1 or 2 fuck-acts/month</td>
<td>75%</td>
<td>73%</td>
</tr>
<tr>
<td>Top 1/10th of individuals account for n% of fuck-acts</td>
<td>34%</td>
<td>46%</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>0.42</td>
<td>0.51</td>
</tr>
</tbody>
</table>

numbers involved (numbers of both acts of and of individuals) and also at the shape of the distributions—the level (average), spread (dispersion), and straggle (concentration at high end and low end). The summary information is given in Table 3.

The most astonishing thing about Table 3 is how similar the four types of highest risk behaviour are to each other, and also how much the characteristics differ from those that hold for all the acts of anal intercourse taken together.

In particular, we found that:

1. The average number of unprotected acts of anal intercourse is considerably smaller than the overall average for all acts of anal intercourse.

2. The overall spread of unprotected distributions is smaller, showing a more compact form and concentration.

3. The low-end concentration (the extent to which the highest-risk anal intercourse is concentrated in the 1-a-month, or at most the 2-a-month, category) is very marked and is a good deal more concentrated at this low level than general anal intercourse is.

4. The high-end concentration (the share of the acts of anal intercourse of the top 10% of individuals) is very considerably lower for the highest-risk acts (about 35% versus 54% for all anal intercourse). So although there are some individuals who have a disproportionate share in acts of unprotected anal intercourse, this share is a lot less than for the overall rate of anal intercourse.

Perhaps more unexpected is the fact that using a condom makes little difference to the characteristics and concentration of the distributions. The differences in high-risk behaviour due to modality is even less marked, except that receptive variants of anal intercourse seem slightly more common. However, this is affected by one extreme outlier who engaged in receptive anal intercourse on an average of twice a day in the month (thereby accounting by his own activity for 14% of the highest-risk sex). The major characteristic of highest-risk acts of anal intercourse, whether protected by a condom or not, is that most of them (well over half) are concentrated in those who only engage in one act per month. With that information, it becomes clear that in terms of number of men involved, highest-risk acts of anal intercourse are concentrated primarily among those who only engage infrequently, and this merits attention.

**Variables Affecting Risk And Its Concentration**

Finally, it is relevant to summarize the results of searching for pockets of “risk-rich” gay men among the sexual diarists. The three most potent differentiating variables are: relationship status (closed/open/no regular relationship); HIV sero-status (tested positive/last test negative/not tested) and concordant-discordant partner status.

**Relationship status.** Closed relationships had a high amount of unprotected AIE, and 14% of those in an ostensibly closed relationship engaged in unprotected anal intercourse at least once in a month with another partner. In open relationships there was less unprotected AIE with a primary partner and more high-risk sex with other partners.

Those with no regular relationship tended to avoid acts of AIE, make more use of condoms, and ejaculate more “on” than “in” a partner. Hence, highest risk AIE is concentrated with the primary partner for the closed relationships, and with other partners for those in an open relationship.

**HIV sero-status.** HIV-positive men have a quite different profile than men testing negative. Positive men had more sex, had more anal sex, and had more anal sex to ejaculation than the negative men — but positive men had considerably more protected AIE, and engaged in a higher degree of receptive AIE than negative men. Those not tested make less use of condoms than either other group.

**Concordant/discordant partner status.** Concordant highest-risk sexual activity (AIE/NC) (whether concordant positive or concordant negative) is far more prevalent than dis-
cordant AIE (30% of sessions are negative concordant and 26% are positive concordant vs. only 2% which are discordant sessions). But sessions where the diastat is negative and where his partner is either untested or his HIV status is unknown tend to involve the highest rates of highest-risk sex: Here, probably, is the greatest reservoir of concern.

DISCUSSION

Behavioural interventions among gay men are predicated upon data findings that anal intercourse among gay men is a minority behaviour, that condom use differs depending on whether the partner is regular or casual, and that condom use is normally consistent with principles of risk-reduction. Most of these findings are broadly supported both by the SIGMA interviews (Davies et al., 1993) and diary studies (Coxon, 1996). If AI is viewed as a necessary condition for risk, then it is clear that it is highly concentrated: There is a significant number of individuals who are responsible for a high fraction of high-risk sex, and their sexual activity can have far-reaching consequences. But the concentration of highest-risk anal intercourse is primarily in the relatively infrequent acts of a relatively large number of gay men (rather than in the very frequent acts of a few)—and it is this which is likely to lead to more rapid diffusion of infection and, ultimately, to higher levels of infected individuals.

REFERENCES


Manuscript accepted June 9, 1999.