National Institute on Drug Abuse

Research 57

MONOGRAPH SERIES

Self-Report Methods of timating ɪg Use

ting Current Challenges lidity

) HUMAN SERVICES • Public Health Service • Alcohol, Drug Abuse, and Mental Health Administration

100257

Self-Report Methods of Estimating Drug Use:

Meeting Current Challenges to Validity

Editors:

Beatrice A. Rouse, Ph.D.
Nicholas J. Kozel, M.S.
Louise G. Richards, Ph.D.
Division of Epidemiology and Statistical Analysis
National Institute on Drug Abuse

NIDA Research Monograph 57
1985

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Alcohol, Drug Abuse, and Mental Health Administration

National Institute on Drug Abuse
5600 Fishers Lane
Rockville, Maryland 20857

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

NIDA Research Monographs are prepared by the research divisions of the National Institute on Drug Abuse and published by its Office of Science. The primary objective of the series is to provide critical reviews of research problem areas and techniques, the content of state-of-the-art conferences, and integrative research reviews. Its dual publication emphasis is rapid and targeted dissemination to the scientific and professional community.

Self-Report Methods of Estimating Drug Use: Meeting Current Challenges to Validity

ACKNOWLEDGMENT

This monograph is based upon papers presented at a technical review on validity issues in self-reported drug use which took place on May 8 and 9, 1984, at Bethesda, Maryland. The meeting was sponsored by the Division of Epidemiology and Statistical Analysis, National Institute on Drug Abuse.

DHHS publication number (ADM)85-1402
Printed 1985

NIDA Research Monographs are indexed in the Index Medicus. They are selectively included in the coverage of American Statistics Index, BioSciences Information Service, Chemical Abstracts, Current Contents, Psychological Abstracts, and Psychopharmacology Abstracts.

PREFACE

In many areas of research, including those sponsored by the National Institute on Drug Abuse, the self-report has become an integral component of the research methodology. While there is a growing body of literature that supports the general veridicality of the self-report, there are also studies that suggest under-reporting in certain populations. This is true of a number of behaviors, including the self-administration of drugs such as tobacco, alcohol, marijuana, cocaine, and heroin. For some behaviors which are highly stigmatized and/or relatively rare events, such as heroin use, the self-report in sample surveys may not be the most appropriate technique. Thus, methods other than self-report have been used to estimate the prevalence of heroin addiction. Self-reporting may vary based on the social acceptance, or perceived acceptance, of the behavior in question. While the level of stigma associated with heroin use has been constant, there have been sharp reversals in the perception of the acceptability of such behaviors as smoking tobacco, marijuana use, and cocaine use. Shifts in society's attitudes toward these behaviors raise legitimate and necessary questions regarding the continued veridicality of the self-report for these particular behaviors, and it then becomes incumbent upon the researcher to address these issues. This monograph is based, in large part, on a technical review that was held to discuss issues of validity of self-reported data as well as various estimation techniques. This meeting was another step in a continuing effort designed to maintain excellence in the Institute's research in this area and to contribute to the field in general.

Edgar H. Adams
Acting Director
Division of Epidemiology
and Statistical Analysis
National Institute on Drug Abuse

CONTENTS

INTRODUCTION AND OVERVIEW

Beatrice A. Rouse, Nicholas J. Kozel
Louise G. Richards

Various methods of identifying drug users have been developed to improve or validate estimates based on direct questioning of individuals regarding their use of drugs. These methods include biochemical analysis of different body fluids, indirect questioning techniques, and statistical modeling procedures. Examples of each of these methods are included in this volume.

Chapters in this volume are the product of a National Institute on Drug Abuse (NIDA) Technical Review convened to examine various methodological issues regarding the validity of self-report data. The meeting was held May 8-9, 1984, at the National Institutes of Health.

While validity covers a wide range of methodological concerns, the Technical Review participants focused on three areas: 1) under-reporting of drug use on direct questioning, 2) noncoverage of groups in the population who are at risk, and 3) procedures for estimating low prevalence drugs, such as heroin. The papers presented in this volume attest to the rich array of concepts, methods, empirical results, and evaluations engendered by the meeting.

In the first chapter, Nurco identifies and describes the different types of validity, raises various validity issues in drug use and crime research, and suggests strategies for improving the validity of self-report data.

Harrell reiterates the fact that validity is a multidimensional concept and discusses those research conditions and respondent characteristics found in methodological research that affect the levels of underreporting. The results of Gfroerer's analysis of privacy support the importance of the conditions under which an interview takes place.

Johnston and O'Malley discuss the advantages and limitations of school surveys in general and describe the Monitoring the Future study in particular. They also address the issue of noncoverage and examine the effect on drug estimates of omitting absentees and dropouts.

Smart presents the Canadian experience with school surveys and suggests some strategies for increasing the cooperation of school boards--a major source of noncoverage in school surveys. In addition, he describes an informant method of deriving drug use estimates.

Shreckengost presents various tests to evaluate the validity of dynamic simulation models and points out that statistical models may often illuminate data that are erroneous.

Frank delineates the methodological issues that affect the validity of telephone surveys and describes the creative ways that the research group at the New York State Division of Substance Abuse Services has dealt with these issues. One of these validation efforts involved the use of a randomized response technique over the telephone.

Zuckerman and colleagues report their results using urine tests with pregnant women to validate self-reported use of marijuana.

Brodsky presents a historical outline of the different techniques, with varying levels of statistical sophistication, used to derive national estimates of heroin users and points out that these methods have changed as both the legal status of opiates and the population subgroups using opiates have changed.

Miller describes the nominative technique, which has been used in the National Institute on Drug Abuse's National Household Survey of Drug Abuse. She presents the logic of the technique and the formula for obtaining a prevalence estimate based on this approach, and compares the rates obtained with those from direct self-report.

Crider compares the trends in incidence of heroin use based on self-report from the National Household Survey of Drug Abuse with such indicator data as hepatitis B cases and heroin-related emergency room visits, deaths, and treatment admissions.

Gardiner and Shreckengost present a dynamic simulation model of the heroin system. They focus on the relationship between supply (inventory) and demand (desired inventory) and how this relationship affects the price of heroin, its purity, the number of heroin-related deaths, and the number of heroin users.

Finally, Woodward, Bonett, and Brecht suggest the multiple-recapture census as a supplemental approach to estimating the prevalence of heroin. They present several mathematical models of this sampling process and the conditions under which each model is most appropriate.

The rich array of methodological issues and techniques presented both during the technical review and in this volume attest to the accomplishments of methodology during the short history of drug abuse epidemiology. Ten years ago, at a similar meeting, the problems were rampant and the solutions sparse. The challenges today also seem numerous, but they are more sophisticated. They require attention to established knowledge as well as creative solutions to meet them. This strategy that evaluates previous experience undergirds all solid development and should lead to even higher quality estimates of drug abuse 10 years from now.

COEDITORS

Beatrice A. Rouse, Ph.D.
Nicholas J. Kozel, M.S.
Louise G. Richards, Ph.D.
Division of Epidemiology
and Statistical Analysis
National Institute on Drug Abuse

A DISCUSSION OF VALIDITY

David N. Nurco

Scientific research rests on a foundation of measurement and classification. The translation or "operationalization" of concepts into indicators or measuring devices and the application of those devices define the process of measurement, without which formal science cannot exist. The term validity enters the researcher's vocabulary when we self-consciously ask ourselves how well we have succeeded in performing the measurements we have undertaken, or when we ask how much trust we can place in the resultant data. Discussion of the validity of estimates of substance abuse in this volume has indicated a role for laboratory procedures in improving the accuracy of self-reports; other discussions have concentrated on our inevitable dependence on self-reports. It seems clear that we are talking about a variety of approaches to validity.

TYPES OF VALIDITY

The concept "validity" is usually defined by the question "Are we measuring what we intend to measure?" Philosophically, this question is either tautological or it is unanswerable. It implies that there is some ultimate reality (perhaps in the mind of God), but that the reality is not available to mere mortals; if it were, we would approach it directly and the validity question would be meaningless. If ultimate truth is out of our reach, then we must be talking about approximate truth (if the logicians will forgive us) and, further, about degrees of approximation. We then must face the fact that we have no objective way to determine which of our approximations is closest to the real thing. Therefore, we must rely upon consensus or common sense to decide, for example, which of two measures is the criterion and which is the candidate for validation.

In dealing with self-reports of substance abuse, we have a further complication. It is fair to assume that objective truth exists. For example, the teenager did or did not use amphetamines. But the validity question is twofold. We use the term "veridicality" to describe the extra link between reality and report. In this age of look-alikes, do respondents really know what they used? In addition, are they telling us the truth as they know it? I might point out in passing that all we can ever hope for in self-reports is that our subjects try to tell us the truth as they believe it exists. We cannot pretend to measure the swindles of look-alikes in surveys.

Validity, which has to do with the accuracy or correctness of what is being measured, must, of course, be differentiated from reliability, which refers to the consistency or reproducibility of measurement, as in repeated measures. If a measurement procedure is not reliable, it will not even agree with itself, so reliability is a necessary but not sufficient precondition for validity. However, the converse is not true; a measure can give consistent (reliable) results and still be invalid, as with a thermometer that always reads 5 degrees too high.

Although reliability and validity are the most often discussed psychometric desiderata, a fuller treatment of the quality of measurement would also refer to: 1) objectivity, or the degree to which the measurement is independent of the person performing it; 2) precision, or the extent to which the measurement is capable of detecting small differences, which may or may not be important; 3) utility, or the adequacy of the measurement for its purpose. For example, the invalid thermometer that always reads 5 degrees too high may be just as useful as a correct one if the purpose is to measure the relationship between temperature variation and discomfort. Much of the art or science of psychometrics deals with ingenious ways to operationalize reliability and validity in order to evaluate the success of the measurement process.

An index of reliability is often calculated by the test-retest (e.g., reinterview) technique, a procedure that yields what is more properly referred to as a stability coefficient of reliability. Any test-retest procedure must deal with such embarrassing considerations as whether the phenomenon itself has changed between measurements or, in the case of cyclic behavior, e.g., hunger, drowsiness, sex drive, whether one is dealing with a different sample of the trait. Reliability has sometimes been taken to mean an agreement among observers, as illustrated by Stephens (1972), who defined reliability as agreement between client and counselor on a number of items. Reliability defined as inter-rater agreement usually refers to situations in which "judges" or equivalent observers are asked to rate or otherwise evaluate certain characteristics. Reliability of scales or tests is frequently measured by internal consistency, indicating the degree to which the separate items making up the test or scale are correlated with one another. Many psychometricians consider this the most important kind of reliability, and it is frequently indexed by Cronbach's coefficient alpha (1970). Split-half reliability is a special case of coefficient alpha, since it measures the agreement between halves of an instrument that has been partitioned in only one of the many ways possible. Coefficient alpha, on the other hand, is the average of all possible split-halves (Cronbach 1970, Nunnally 1978).

When we get into the area of validity, the picture becomes even more complex. Although distinctions are sometimes blurred, psychometricians approach validity in a variety of ways. Face validity means that, in some expert's judgment, an instrument, questionnaire, or other device appears to measure what it purports to measure. Content validity goes further in that the instrument consists of items that explicitly deal with the issue at hand, as in a vocabulary test that consists of words to define. Concurrent validity refers to the degree of agreement between the test results and some other measure of the same thing that is obtained (approximately) concurrently and that is generally regarded as valid on a priori grounds. Predictive validity is much the same as concurrent validity except that the behavior or event to be predicted, termed the criterion and believed to be a better indicator of the phenomenon in question, will occur sometime in the future. Finally, and perhaps most important for theoretical purposes, there is construct validity. Although this is somewhat difficult to define, it is closely linked with factor analytic strategies and the notion of a latent variable. A latent variable (or its equivalent, a hypothetical construct) has no single, pure criterion or indicator of itself. Rather, it is an abstraction measured by the relationships among a number of observable variables believed to be partially determined by it. Some examples of latent variables that are widely used in the social and behavioral sciences are intelligence, anxiety, depression, socioeconomic status, and gross national product. In effect, these latent variables are measured by obtaining a weighted combination of the several indicators believed to be determined by them. Note that it is impossible to directly validate such a measure because no direct criterion exists. However, if the results of our measure are in accord with theoretical expectation, e.g., if persons under a psychiatrist's care score higher on a measure of anxiety than do people not under such care, we conclude that the measure is behaving in a manner consistent with theory.

Turning more directly to the issues that are of present concern, we might begin by saying that reliability is only a means to an end--validity--and that validity is a means to an end--utility. The real question is "How are you going to use the valid measure?" For example, the precision of each individual's response is important if the objective is individual diagnosis, but a considerable amount of imprecision may be tolerable over a sizable sample of individuals if the objective is the study of correlates of a response. Even if the objective involves determination of a prevalence estimate for a population, and policy is being made on the basis of the estimate, only policymakers can tell you whether they would behave differently if there were 1 million heroin addicts than if there were 900,000.

To illustrate the difference between individual diagnosis and group findings, consider evaluation of drug use among Vietnam veterans (Nurco, unpublished data). Based on self-reported data, 10 percent of the subjects admitted drug abuse; laboratory reports of these same subjects showed that 10 percent had positive urine tests. However the admitted drug users were not all the same 10 percent of the subjects with the positive urine tests.

Sometimes, we are tempted to speak glibly of comparing self-reported drug use against a criterion of medical records or other ostensibly superior sources. We tend to forget that what appears in the medical records is what the subject told the recordkeeper. In effect, we are merely comparing the results of one interview with the results of another interview. We must remember that, at best, we have highly fallible criteria, such as clinic records of uneven quality, ancillary reports of family and friends, police records of varying degrees of completeness and accuracy, counselor appraisals of questionable reliability and objectivity, and inconsistent laboratory results.

VALIDITY ISSUES IN DRUG USE AND CRIME RESEARCH

With regard to more objective criteria of legal involvement, the deficiencies inherent in official criminal records have been thoroughly reviewed and documented by Collins and his associates (1982). In their summary regarding various aspects of data quality, they state:

> There is considerable evidence in the literature that individuals, when asked about arrests, attempt to report that information accurately. It is also clear from the literature that individuals sometimes do not report their involvement in crime and that the accuracy of arrest self-reports varies on the basis of length of recall period, type of criminal behavior, and data collection methodology. It is not possible to summarize the nature of the systematic bias that exists in individual reporting patterns in any simple way. The evidence is not consistent.
>
> It is also clear from the literature review that official records are often deficient. Arrests are not always entered on an individual's record, and the accuracy of a given record depends on the types of offenses, where the arrest took place, and when the arrest occurred. In the past there has been a tendency for researchers to be concerned about the validity of self-reports and to ignore the deficiencies in official crime records. The preceding review clearly shows that official records should not be viewed as complete and accurate (Collins et al. 1982, pp. 16-17).

An issue not raised by Collins et al. but sometimes alleged is that official reports of crimes may be subject to political pressures to show a decrease in certain types of criminal activity. If this is true, there is even further reason to question the value of official records as a criterion against which to validate information from other sources.

On a more optimistic note, I want to emphasize and elaborate on some points made in Harrell's paper in this volume. For those of us in narcotic addiction research, it is encouraging to encounter the view that addicts might indeed be more honest respondents regarding drug abuse than others in the general population by the very fact that they have little to conceal. The addicts that we deal with in our studies are usually well known to authorities and have already been involved in rehabilitative efforts.

Also, admitting to addiction may have its compensations. Though generally we must be concerned about underreporting, overreporting may be a problem in some instances. For example, individuals arrested for violent crimes, with the prospect of prolonged incarceration, may present themselves as victims of narcotic dependence--not entirely responsible for their actions and more in need of treatment than punishment. Lack of veridicality in the form of overreporting, therefore, may present more of a problem to judicial and correctional personnel than to survey researchers in the community. Overreporting is also an issue that has to be dealt with in any desirable or popular treatment program. For example, overreporting tends to be a problem in determining eligibility for methadone maintenance programs and has given rise to the use of narcotic antagonist challenges to confirm addiction. (This raises the problem of frequency of use versus actual physical dependence in determining addiction. Addicts may be truthfully reporting extent of use but may not actually be addicted.)

IMPROVING VALIDITY OF SELF-REPORTED DATA

We should not, however, give up on self-report data. Rather, we should concentrate on making it better. From information presented by Harrell (this volume), we can devise measures to reduce concealment or underreporting.

Among strategies to be considered are:

1) Assuring confidentiality of information

2) Establishing rapport
    a. Selecting empathic and skillful interviewers
    b. Enlisting respondent support by presenting general objectives of the study, e.g., appeal to altruism

3) Checking records and informing subject of intent, which should be beneficial not only as a concurrent check but may actually improve accuracy of self-report

4) Urine monitoring and informing subject of this intent, which should be beneficial not only as a concurrent check but may actually improve accuracy of self-report

5) Concentrating on recent events

6) Making questions less specific

Dr. Harrell's suggestion that researchers check questionnaires for bias during test construction is worth repeating: the preparation of any study using self-reported data should include pretesting of the questionnaire and field work procedures to evaluate the potential response bias associated with the mode of inquiry. This fits into the widely followed practice of tailoring questionnaires to a specific purpose and of careful test construction as part of most research conducted in the area of drug abuse. The notions seem to be strongly supported by Mr. Gfroerer's analysis of responses under varying conditions of privacy, since privacy in this instance is part of the assurance of confidentiality. As Mr. Gfroerer points out, his findings are clouded by the fact that the level of privacy in each interview was not randomly assigned, so that unsuspected exogenous variables may have exerted an effect.

Dr. Zuckerman (this volume) discusses the validity of reported marijuana use as determined by concurrent laboratory findings. Researchers also should be concerned about the validity of reports of alcohol use and smoking. Both of these activities are extremely important with respect to pregnancy, and it is imperative that we learn more about them. Laboratory tests for these substances have improved and are available; therefore, the validity of reports of alcohol and tobacco use as well as marijuana should be checked. In fact, an intriguing design would be to test the hypothesis that validity of self-report diminishes with increased social undesirability of reported activity, e.g., use of cigarettes, alcohol, and marijuana.

Researchers' main concern has been obtaining valid results, but consideration of this objective prompts a related concern: granted the validity of the findings, how valid is their generalization to the population in question? Or, put another way, are our results valid for only a subsection of our population? This raises the issue, discussed earlier, of utility as the ultimate objective of validity.

To illustrate this point, consider our experiences in our natural history study (Nurco 1975, Nurco et al. 1981a, 1981b, 1981c), which involved the examination of addict careers. We used a roster of individuals already identified in the arrest and investigation files of the Baltimore City Police Department's Narcotic Squad. Since we were planning to study narcotic addicts, we were concerned with how representative our sample was. With this in mind, we identified a number of addicts in the State mental hospitals in Maryland who were not then known to the police. When we checked, 3 years later, we found that virtually all these appeared on an updated police roster. The fact that all addicts eventually became known to the police helped to calm our fears about the generalizability (and thus the utility) of the original sample.

A full consideration of representativeness not only involves sample identification but is also concerned with accessibility and subsequent attrition. In the first wave of our study, we selected 10 white males and 10 black males from those newly identified during each year of the period 1952 through 1966, and 5 white males and 5 black males for the period 1967 through 1971. (We oversampled in the earlier years because of our interest in the careers of addicts). We located 98 percent of our sample. Our interview response rate was in excess of 92 percent, with an interview that took approximately 3 hours to administer. Unfortunately, the success of this endeavor in producing this level of response represents the exception rather than the rule in drug abuse research. I present it here to emphasize the importance of obtaining a high response rate and as an example of what can be achieved through perseverance.

With regard to the veridicality of the information obtained, we wanted detailed data about drug use, employment, criminal behavior, and social relationships for each period of addiction over a lifetime involvement with drugs. In a pilot phase, we found that the typical subject would collapse his periods of addiction with periods of nonaddiction as a way of bringing a long interview to conclusion. As a result of this experience, our subsequent strategy was to determine the dates of successive addiction periods before asking detailed questions about each one. After eliciting these dates, we asked questions about preaddictive behavior and then moved to each of the on-and-off periods we were interested in, reminding the subject of the dates we had originally elicited from him. In this way we obtained the information we needed.

I am not suggesting that a tactic similar to the above be used in every study. I am recommending, however, that researchers in the field steep themselves in the nuances of veridicality until they appreciate the magnitude of the problem and are prepared to devise anticipatory strategies to avoid its many pitfalls.

REFERENCES

Collins, J.J.; Hubbard, R.L.; Rachal, J.V.; Cavanaugh, E.R; Craddock, S.G.; and Kristiansen, P.L. Criminality in a Drug Treatment Sample: Measurement Issues and Initial Findings. Report to the National Institute on Drug Abuse under Contract 271-79-3611. Research Triangle Park, N.C.: Research Triangle Institute, 1982.

Cronbach, L.J. Essentials of Psychological Testing. 3rd ed. New York: Harper and Row, 1970.

Nunnally, J.C. Psychometric Theory. 2nd ed. New York: McGraw-Hill, 1978.

Nurco, D.N.; Bonito, A.J.; Lerner, M.; and Balter, M.B. Studying addicts over time: Methodology and preliminary findings. Am J Drug and Alcohol Abuse, 2:183-196, 1975.
Nurco, D.N.; Cisin, I.H.; and Balter, M.B. Addict careers. I. A new typology. Inter J Addict, 16:1305-1325, 1981a.
Nurco, D.N.; Cisin, I.H.; and Balter, M.B. Addict careers. II. The first ten years. Inter J Addict, 16:1327-1356, 1981b.
Nurco, D.N.; Cisin, I.H.; and Balter, M.B. Addict careers. III. Trends across time. Inter J Addict, 16:1357-1372, 1981c.
Stephens, R. The truthfulness of addict respondents in research projects. Inter J Addict, 7:549-558, 1972.

AUTHOR

David N. Nurco, D.S.W.
Department of Psychiatry
University of Maryland
1229 West Mt. Royal Avenue
Baltimore, MD 21217

VALIDATION OF SELF-REPORT: THE RESEARCH RECORD

Adele V. Harrell

Self-reported data are the mainstay of much social research. Indeed, questionnaires and checklists have become a way of life as we record our food preferences or television viewing for the latest survey. The popularity of self-reported data can be easily understood. It is relatively simple to collect--by mail, by telephone, by face-to-face interview, or with self-administered questionnaires. The ability to manipulate the mode of questioning and the questionnaire content provides a great deal of flexibility in designing studies. In addition, certain types of information can be collected from individuals with less effort and often more accurately than from alternative data sources. Take, for example, questions such as "How old were you when you got a driver's license?" or "How often have you been hospitalized?" Searches of the records either at the Division of Motor Vehicles or at hospitals would be far more time-consuming than a questionnaire and would depend heavily on the accuracy of the officially maintained records. More important, there is certain information that can come only from the individual. This includes, for example, information on private personal behavior, such as voting behavior, and information about individual attitudes. Small wonder that we rely so heavily on easy-to-get self-reported data. However, in the face of the good news about self-reported data, it is necessary to take some time to consider the bad news--or at least the potential for bad news.

The potential for bad news comes in the form of multiple threats to the validity of self-reported data. Validity in this context refers to whether the data recorded by the researcher accurately reflect the phenomenon under investigation. This simple statement conceals what is in actuality a complex, multidimensional concept. Validity can take on a variety of meanings, depending in part on the method used to evaluate the extent to which the data reflect the phenomenon under investigation. Face validity, for example, refers to the extent to which the data appear to "make sense" as a reasonable indicator of the purported phenomenon. Predictive validity refers to the extent to which the data correlate with subsequent outcomes to which they should be related on logical grounds. Criterion validity, the primary focus of this discussion of self-reported data, refers to the extent to which the subjective self-reported data are "verified" by agreement with another indicator of the same phenomenon believed to be of higher validity (the criterion). Criteria used for this purpose have included official or medical records; reports of others such as family, friends, or counselors; and biochemical tests (Stanton 1972).

Any number of factors can undermine the validity of self-reported data. These include careless field procedures (Deming 1950), question design and content (Bradburn et al. 1979), memory lapse (Deming 1950), and status bias (Cahalan 1968). In general, the validity threats cited by researchers fall into three categories:

1) Aspects of the mode of inquiry; factors in the questioning situation that influence the response. Examples include question wording, interviewer expectations, and degree of anonymity.

2) Inability to provide correct information; respondent never knew or has forgotten the answer and thus cannot provide valid data.

3) Unwillingness to provide the information; respondents' answers are designed to present them in a socially favorable way and/or to promote their personal interests. In this case, the respondent is unwilling to provide information requested.

Of these threats to validity, it is the third category that is a crucial issue in studies of illicit drug use. Illicit drug use is behavior that carries with it the threat of social sanctions and the stigma of illegality. The negative social status of illicit drug use may deter some survey respondents from accurately reporting their drug use experiences--either in an attempt to avoid adverse reactions from parents, employers, or teachers (if not peers) or in an attempt to present themselves in a favorable way during the interview. This concern is not without theoretical foundation. Social desirability theory (Edwards 1957) rests on the premise that the more highly stigmatized and negatively sanctioned a behavior, the stronger the tendency to deny having engaged in it. This theoretical perspective indicates that distorted responses, either underreporting or overreporting, will occur as a function of the perceived acceptability of the correct response. The following review, which begins with a general look at the validity of self-reported data and goes on to examine drug use validity studies, provides empirical evidence consistent with this thesis.

The focus on veridicality as a central issue is not new. Hyman aptly entitled his 1944 article "Do they tell the truth?" His subject was the socially sensitive issue of the redemption of war bonds. In the midst of World War II, cashing in war bonds was widely thought to be unpatriotic and the basis for strong social censure. Hyman found that a substantial percentage of persons known to have cashed in war bonds denied having done so within a week of the redemption. Apparently, respondents were unwilling to admit such socially unacceptable behavior.

The importance of the apparent "degree of deviance" is illustrated more clearly in the Denver validity study (Parry and Crossley 1950), which compares the accuracy of response for socially neutral and socially positive behaviors. Respondents were asked about their voting record, charitable contributions, possession of a library card, possession of a driver's license, and ownership of specific items such as a telephone. Records from the polls, the Community Chest, the library, etc., were used as the criteria.

Rates of distortion appeared clearly linked to the desirability of the correct response and/or the ease with which the correct response could be verified. Easily verified questions of fact--e.g., telephone ownership, age--attained accuracy rates of 90 percent or higher against external criteria. In contrast, questions about socially desirable behavior produced response distortion. For example, 34 percent of the respondents reported that they had made contributions to the Community Chest that were not recorded in Community Chest records. Further evidence of exaggerated responses about socially positive behavior is found in the overreporting of voting: 16 percent inaccurately reported voting in a recent election, while 42 percent overreported voting in the past six elections. Ownership of a library card, a more neutral behavior, was apparently exaggerated by only 9 percent of the sample. Interestingly, the rate of error in the other direction--failure to report a real contribution to the Community Chest or a vote--was consistently small (under 5 percent) and probably should be attributed to memory lapse, field procedure problems, and other sources of unreliability.

Research on the validity of data on socially undesirable behavior produced similar results. Cannell and Fowler (1963) compared self-reported data on hospitalization to data from records. They found that the denial of hospitalization for threatening or embarrassing disorders was considerably higher than the denial of hospitalization for other kinds of disorders. Phillips and Clancy (1970) found that the willingness to report symptoms of mental illness was directly related to the respondent's view of the social undesirability of various psychological disorders.

The tendency to exaggerate socially positive behavior and to underreport socially negative behavior does not appear to have diminished with time. In 1979, 29 years after the Denver validity study, Bradburn et al. again report substantial distortion of voting behavior--rates of overreporting in response to whether the person voted in a recent primary ranged from 36 percent to 48 percent depending on the mode of inquiry. Drunken driving, with its socially negative status, elicited respondent denial at rates that ranged from 35 percent to 54 percent, depending on the questioning technique. As before, a much lower rate of overreporting characterized responses to the more neutral question of whether the respondent had a library card. These highlights of the literature on self-reported data demonstrate that in general there is a tendency for respondents to give answers that make them "look good."

Not everyone, it appears, is equally likely to give a distorted response. Hyman (1944), for example, found that the rate of denial for war bond redemption increased with income; that is, persons with higher incomes were more reluctant to admit to this behavior than persons with lower incomes. Cahalan (1968) reported that the distorted responses to the Denver validity study varied by sex, age, and socioeconomic status. Compared to men, women were slightly less likely to exaggerate their voting record, and noticeably less likely to exaggerate their Community Chest contributions and possession of a valid driver's license. Younger persons were more likely than older persons to overreport socially desirable behaviors such as voting and contributing. In addition, contributions were more likely to be exaggerated by lower socioeconomic status respondents than by higher status respondents. A similar conclusion was reached by Weiss (1968). She examined self-reported registration and voting among black welfare mothers and found that rates of voting exaggeration were related to age, education, and social status (as well as rapport with and social distance from the interviewer). These exaggerated reports belonged to women she refers to as the "almost voters"--older, more educated women who were more experienced in the labor market and who held middle-class views. They were, in short, women who valued voting and thus tended to present a view of themselves that included voting.

These sociodemographic differences in validity of self-reported data appear in many instances to be a function of personal norms and self-expectations. For example, we can speculate that having a driver's license was, at the time of the Denver Validity Study, less common for women than men. Thus, women may have responded on the basis that having a driver's license was not socially expected behavior--not the norm. Similarly, the wealthy may have held more stringent expectations of their financial responsibilities for the war effort and, therefore, responded as they felt they should have behaved. Weiss clearly believes the personal value system of her sample of mothers influenced the validity of their responses. This interpretation is supported by one study of the validity of self-reported data on deviant behavior (Clark and Tifft 1966). Using anonymous questionnaires, interviews, and polygraph tests to gather information on behavior and attitudes, Clark and Tifft concluded that response inaccuracy was highly related to declared personal norms and reference group norms and was related as well to the generally understood "deviance" of the behavior. Such interpretations are consistent with social desirability theory.

These findings are significant for the study of illicit drug use--a phenomenon that has sharply divided our society into segments characterized by divergent drug use norms both in terms of what is typical or expected behavior and in terms of what is desirable behavior. Thus, the validity of self-reported data on illicit drug use may well vary by age, location, or other correlates of these normative differences.

Much less information is available on the second threat to validity--the respondent's ability to provide the information. In one test of response validity, Jeager and Pennock (1961) interviewed appliance owners twice, at a 1-year interval, on the kind and condition of their washing machines and the year of purchase. The results indicate widespread inability to recall detailed information such as the year of purchase. Almost 60 percent of the respondents contradicted themselves on the year of purchase, although 75 percent of the time the difference was 2 years or less. Agreement on kind and condition of the machine was much higher, illustrating the higher validity for easier questions. This provides clear warning about attempting to collect detailed data without adequate memory aids and carefully designed questioning procedures. However, it is not clear that these results provide any more than general guidelines for the study of illicit drug use. Memory is a selective process, known to be affected by the salience of an event. Events as different as purchasing a washing machine and using cocaine for the first time may well differ in salience, with unknown consequences for the validity of self-reported data.

A third threat to validity in the list was the mode of inquiry. The literature abounds with evidence on how relatively subtle differences in the context of an interview can influence the responses. A classic example was reported in early surveys of drinking, in which interviewers who drank elicited reports of more drinking among respondents than did interviewers who abstained (Mulford and Miller 1959, 1963; Mulford 1964). Similarly, Weiss (1968) reported that interviewer rapport with her sample of welfare mothers influenced the validity of their self-reported voting behavior.

Variation in the mode of questioning can affect the degree to which responses are anonymous and thus the degree of self-disclosure required (Jourard 1971). The work of Bradburn et al. (1979) is particularly interesting because it relates variations in degree of anonymity to the validity of responses to questions about behaviors that vary in social desirability. They evaluated the responses to questions ranging from drunk driving to voting collected by telephone, by mail, by face-to-face interview, or with self-administered questionnaire. Although theoretically one might expect more authentic responses under conditions of greater anonymity, the reported validity differences due to variation in questioning techniques are so small as to have little practical import, regardless of the behavior's social desirability.

This general overview provides the background for evaluating the relatively sparse literature on the validity of self-reported data on illicit drug use.

Much of the research on the validity of self-reported data on drug use has focused on the veridicality of former narcotic addicts. These studies compare addicts' reported drug use, arrest record, and demographic information to data from hospital records, law enforcement records, biochemical tests, and/or reports of significant others (Ball 1967; Cottrell and O'Donnell 1967; Robins and Murphy 1967; Stephens 1972; Amsel et al. 1976; Maddux and Desmond 1975; Bonito et al. 1976). In one such study, Ball (1967) compared the responses of 59 narcotic drug addicts in a structured interview to data from hospital records, FBI records, and urine tests conducted immediately after the interview. His expressed goal was to determine whether "deviant groups, especially those engaged in illegal behavior, are motivated to--and do--conceal or deny their proscribed behavior" (Ball 1967, p. 650). Five items were used for comparison: 1) the age of the subject, 2) age at onset of drug use, 3) type and place of first arrest, 4) total number of arrests, and 5) drug use at the time of interview. Responses to the items related to deviant behavior "indicate a rather surprising veracity on the part of former addicts" (Ball 1967, p. 653). In fact, the only response bias noted was that the females in the sample were significantly less reliable than the males in reporting their age. In general, most research on former addicts concludes that addicts are willing to reveal the facts of their drug use and arrest record, although Amsel et al. (1976) report relatively high denial rates for drug use.

Recall of detailed information does, however, appear to pose a threat to validity for some drug use items. Higher rates of distortion are reported for exact information, e.g., age at first arrest and age of first drug use (Cottrell and O'Donnell 1967, Ball 1967) than for "easier" questions, such as "Have you used marijuana?" Because the addicts appeared willing to provide authentic drug information, the implication is that faulty memory produces these inaccurate answers.

Although these results are quite encouraging for those who wish to gather drug use data in interviews, the generalizability of the addict studies to surveys of the general household population is questionable. Previously hospitalized addicts have already been publicly labeled deviants and identified as drug users. They are well aware that records of their drug-use history exist. This reduces the amount of new self-disclosure required. The use of urinalysis in some studies further discourages attempts at concealment by reducing the chances of successful concealment (Amsel et al. 1976). In contrast, nonaddict drug users in the household population may be actively engaged in concealing their use from others and may believe the drug-use behavior to be unknown except to selected persons. In these cases, there may be greater incentives for denial and greater chances of successful concealment.

One study of the validity of self-reported data conducted with members of the household population was reported by Parry et al. (1971). The research evaluated the accuracy of responses to questions on the use of psychotherapeutic prescription drugs, using as a criterion a complete file of all prescriptions filled in a small midwestern town. Antibiotic use, believed to be less stigmatized than use of tranquilizers, stimulants, or sedatives, was used for comparison. The overall level of response accuracy can only be described as fairly good: the percentage of respondents who correctly reported their drug use ranged from 64 percent to 80 percent. However, the accuracy appears to be unrelated to the social desirability of the drug. Indeed, the accuracy rate for psychotherapeutic drug use averaged 74 percent, exceeding the accuracy rate of 64 percent for antibiotics. The tendency of adults in the household population to underreport psychotherapeutic drug use is also reported by Fejer and Smart (1973), who found that the rate of psychotherapeutic drug use calculated from pharmacy dispensing records was substantially higher than the rate estimated from self-reported survey data.

Once again, respondent inability to give accurate data may have posed a problem. The use of a long-form questionnaire plus drug charts in color by Parry et al. (1971) reduced the percentage of incorrect responses on tranquilizer use from 27 percent to 15 percent. The effectiveness of these techniques in reducing incorrect answers suggests that respondents will provide valid answers if questions are presented in a way that facilitates recall. The study further indicates the effect of time on the accuracy of the responses; inaccurate responses were twice as prevalent among former tranquilizer users as among current users. This is a sharp decline in view of the relatively narrow time span used to define former users (used in 1966 or 1967, but not in 1968 or 1969) and current users (used in 1968 or 1969). The results thus indicate a marked drop in accurate response within a year or two of the drug use. It should be noted, however, that this study differs in one important aspect from surveys of illicit drug use. While the medical use of the psychotherapeutic drugs may carry a negative social image, it would generally be viewed less negatively than illicit use of drugs such as cocaine or heroin.

Of all the studies reviewed, the one that comes closest to addressing the validity of self-reported drug-use data collected in a household interview was conducted in San Francisco just prior to the first National Survey on Drug Abuse (Cisin and Parry 1980). The design compared responses of patients at selected drug abuse clinics with responses of a matched control group and with the clinic records. Questions covered nonmedical use of psychotherapeutic drugs; use of marijuana or hashish; and use of heroin, other opiates, hallucinogens, or cocaine. For most drugs the questionnaire data revealed at least as much use as the clinic data, though both types of data were subject to response errors. However, for heroin, the clinic records revealed considerably more use. For example, the questionnaires detected 73 percent of the cocaine use reported to the clinic, while the clinic records listed 60 percent of the cocaine use reported on the questionnaires. In contrast, the questionnaires detected only 41 percent of the heroin use reported to the clinic, while the clinic records listed 85 percent of the heroin use reported on the questionnaires.

Thus, it seems that the answer to Hyman's question, "Do they tell the truth?" is: not always. There is a clear tendency for respondents to exaggerate the extent of their socially desirable behavior and minimize or deny their socially undesirable behavior. However, beliefs as to what constitutes socially desirable or undesirable behavior seem to vary widely. Differences in personal values, expectations, and reference group norms appear to be key factors in how willing respondents are to provide authentic responses. In the area of drug use, ex-addicts appear relatively truthful--willing to admit candidly their arrest record and drug experience. However, household interviews with drug clinic patients, a group less "deviant" than previously hospitalized addicts, revealed a reluctance to admit to the use of heroin. As Bonito et al. (1976) put it, "Despite the consistently positive evidence for addict veridicality, we believe it important that every study relying on self-disclosure provide evidence for the essential accuracy of the information obtained" (p. 720).

A second finding is that there are definitely limits to the kind of detailed information respondents are able to recall. Certainly, recent events are more accurately recalled than past events. Likewise, precise information is more difficult to recall than general information. That is, use of marijuana is easier to recall than the number of times marijuana was used. Some types of information, e.g., the name of the particular tranquilizer used, require the use of memory aids to achieve accuracy. In this regard, it is wise to consider carefully the degree of precision that is actually required for research purposes. In most, if not all, cases, it is far better to get valid answers to general questions than biased answers or nonresponse.

The ability of a survey to elicit truthful answers based on accurate recall may be a function in part of the mode of inquiry, of the situation in which the respondent is asked to reveal personal information. The cumulative evidence suggests that bias related to the mode of inquiry may vary from one research setting to another. That is, anonymity may be a necessary component in some situations, while interviewer characteristics may be important in others. It is safe to say that the preparation for any study using self-reported data should include pretesting of the questionnaire and field work procedures to evaluate the potential response bias associated with the mode of inquiry.

Although these potential threats to the validity of self-reported data do indeed require thought and preparation on the part of any researcher, they by no means should be construed as indications that self-reports should be abandoned. Self-reported data remains a flexible and relatively efficient method of gathering information that might otherwise be inaccessible. The message is that we must be constantly vigilant and ready to examine such data with a critical eye.

REFERENCES

Amsel, Z.; Mandell, D.; Matthias, C.; et al. Reliability and validity of self-reported illegal activities and drug use collected from narcotic addicts. Int J Addict, 11(2):325-336, 1976.

Ball, J. The reliability and validity of interview data from 59 narcotic drug addicts. AJS, 72:650-654, 1967.

Bonito, A.J.; Nurco, D.N.; and Shaffer, J.W. The veridicality of addicts' self reports in social research. Int J Addict, 11(5):719-724, 1976.

Bradburn, N.M.; Sudman, S.; et al. Improving Interview Methods and Questionnaire Design. Washington, D.C.: Jossey-Bass, 1979. 214 pp.

Cahalan, D. Correlates of respondent accuracy in the Denver validity survey. Public Opin Q, 32:607-621, 1968.

Cannell, C.F., and Fowler, F.J. Comparison of a self-enumerative procedure and a personal interview: A validity study. Public Opin Q, 27:250-264, 1963.

Cisin, I.H., and Parry, H.L. Sensitivity of survey techniques in measuring illicit drug use. In: Rittenhouse, J.D., ed. Developmental Papers: Attempts to Improve the Measurement of Heroin in the National Survey. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1980. pp. 3-46.

Clark, J.P., and Tifft, L.L. Polygraph and interview validation of self-reported deviant behavior. Am Sociol Rev, 31:516-523, 1966.

Cottrell, E.S., and O'Donnell, J.A. Reliability of admission data. Unpublished report. National Institute of Mental Health Clinical Research Center, Lexington, KY, 1967.

Deming, W.E. Some Theory of Sampling. New York: Wiley, 1950.

Edwards, A.L. The Social Desirability Variable in Personality Assessment and Research. New York: Dryden, 1957. 602 pp.

Fejer, D., and Smart, R.G. The use of psychoactive drugs by adults. Can J Psychiatry, 18:313-320, 1973.

Hyman, H. Do they tell the truth? Public Opin Q, 8:557-559, 1944.

Jeager, C., and Pennock, J.L. An analysis of consistency of response in household surveys. J Am Stat Assoc, 56:320-327, 1961.

Jourard, S.M. Self-disclosure: An Experimental Analysis of the Transparent Self. New York: Wiley, 1971. 248 pp.

Maddux, J.F., and Desmond, D.P. Reliability and validity of information from chronic heroin users. J Psychiatr Res, 12(2):87-95, 1975.

Mulford, H.A. Drinking and deviant drinking, U.S.A., 1963. Q J Stud Alc, 25:634-650, 1964.

Mulford, H.A., and Miller, D.E. Drinking in Iowa: I. Socio-cultural distribution of drinkers. Q J Stud Alc, 20:704-726, 1959.
Mulford, H.A., and Miller, D.E. The prevalence and extent of drinking in Iowa, 1961: A replication and an evaluation of methods. Q J Stud Alc, 24:39-53, 1963.
Parry, H.J., and Crossley, H.M. Validity of response to survey questions. Public Opin Q, 14:61-80, 1950.
Parry, H.J.; Balter, M.B.; and Cisin, I.H. Primary levels of under-reporting psychotherapeutic drug use. Public Opin Q, 34:582-592, 1971.
Phillips, D.L. Knowledge From What? Theory and Methods in Social Research. Chicago: Rand-McNally, 1972. 204 pp.
Phillips, D.L., and Clancy, K.J. Response bias in field studies of mental illness. Am Sociol Rev, 35:503-515, 1970.
Robins, L.N., and Murphy, D.E. Drug use in normal population of young Negro men. Am J Public Health, 57:1580, 1967.
Smart, R. Percent studies of the validity and reliability of self-reported drug use, 1970-1974. Can J Criminol Corr, 16(4):326-333, 1974.
Stephens, R. The truthfulness of addict respondents in research projects. Int J Addict, 7(3):549-558, 1972.
Weiss, C. Validity of welfare mothers' interview responses. Public Opin Q, 32:622-633, 1968.

AUTHOR

Adele V. Harrell, Ph.D.
Institute for Social Analysis
1625 "K" Street
Washington, DC 20006

INFLUENCE OF PRIVACY ON SELF-REPORTED DRUG USE BY YOUTHS

Joseph Gfroerer

INTRODUCTION

Underreporting of drug use by survey respondents has always been a major concern of drug abuse survey researchers. This concern particularly applies when the respondents are youths since they might fear being punished if their use of drugs was discovered by parents. A number of studies have been conducted to determine whether underreporting is a serious problem and to identify procedures that can be used to obtain the most valid data. While some of the results have been contradictory, most studies conclude that reliable, valid self-reported drug use data can be obtained (Smart and Jarvis 1981; O'Malley et al. 1983; Hubbard et al. 1976; Single et al. 1975; Smart 1975).

Factors that have been identified as possibly affecting the reporting of drug use by youths include the type of questionnaire (interview vs. self-administered), characteristics of the interviewer, the degree of anonymity of the respondent, the setting (home vs. school), and the degree of privacy during the interview (Johnston, in press; Sudman and Bradburn 1974). Degree of privacy refers to the presence of a parent or other person in the same room during the interview.

In several studies, self-administered questionnaires have been shown to produce higher reported prevalence of drug use than interviews. In one study (Hochstim 1967), women were found more likely to report having used alcohol on a self-administered questionnaire than in a face-to-face interview. A study of college students (Krohn et al. 1975) showed higher prevalence of various illegal behaviors such as marijuana use and drinking under age reported on a self-administered questionnaire, although the sample size for the study was very small and differences were not statistically significant. This study also showed that respondents are more likely to report illegal behavior to "hip" interviewers than to "straight" interviewers.

Anonymity and setting have not been clearly shown to be important factors affecting underreporting. King (1970) compared drug use data from identifiable and anonymous questionnaires filled out by recent college graduates and found no significant differences in drug use prevalence between the two approaches. A study of undergraduate students (Luetgert and Armstrong 1973) found similar results. This study also found that with less anonymity, current marijuana users may report their use as being in the past rather than current, unless they are frequent users. A study of youths involved completing a drug and health questionnaire both at school and at home, with parents also completing the questionnaire in a different room (Needle et al. 1983). Results indicated that the setting in which respondents complete questionnaires is not related to reporting bias. Similar results were found in another study (Zanes and Matsoukas 1979) using matched samples of youths. This study also concluded that the presence of parents in the room during the completion of a self-administered questionnaire had no effect on the reporting of youth drug use.

While many of the factors that may affect underreporting can be controlled by researchers, it is not possible to achieve complete privacy (i.e., no other person in the room) in every interview when conducting a household survey. Given this limitation, it is important to assess the impact of the lack of privacy on the results of a survey, both to assess the potential impact on the validity of data from that survey and also to provide general information on the importance of privacy for future surveys. Only one previous study is known to address this specific issue for self-administered questionnaires, and it was based only on a sample of eleventh graders in one school located in a stable, mostly white, middle- and upper-middle-class neighborhood (Zanes and Matsoukas 1979). The present study attempts to further examine this issue using data from a national probability sample of youths 12 to 17 years of age in households.

METHOD

Data from the 1979 (Fishburne et al. 1980) and 1982 (Miller et al. 1983) National Survey on Drug Abuse were analyzed. In these household interview surveys, data were collected from nationally representative samples of youths ages 12 to 17 years and adults aged 18 years and older. Depending on household composition, interviews were conducted with one adult only, one youth only, or both an adult and a youth. Sample sizes for youths were 2,165 in 1979 and 1,581 in 1982. The surveys collected data on whether respondents had used various licit and illicit drugs in the past month (monthly use), in the past year (annual use), or ever (lifetime use).

Interviewers tried to ensure privacy by requesting that the interview be conducted in a private room. Also, most drug-use questions were filled out on an answer sheet by the respondent, so no verbal responses could be overheard by other persons in the house.

Furthermore, confidentiality was promised by having the respondent place completed answer sheets in an envelope and sealing the envelope, assuring the respondent that the interviewer would never see the responses. After each youth interview, the interviewer filled in a question regarding the degree of privacy during questioning. Interviewers rated privacy on a scale of 1 to 9 as follows:

1 - completely private

3 - minor distractions

5 - parent in room around one-third of the time

7 - serious interruptions of privacy more than half of the time

9 - constant presence of parent

Values of 2, 4, 6 or 8 could also be used by interviewers. Respondents with unknown privacy were excluded from this analysis, resulting in sample sizes of 2,148 12- to 17-year-olds in 1979 and 1,538 12- to 17-year-olds in 1982.

The relationship between privacy and reported lifetime use of several drugs was analyzed using weighted linear regression analysis (Draper and Smith 1966). Lifetime use was chosen because of the very small proportion of youths with annual or monthly use of some drugs and because underreporting is believed to be less serious for lifetime use than for annual or monthly use (Luetgert and Armstrong 1973). Thus, if lack of privacy is found to affect reporting of lifetime use, reporting of current use is probably affected at least as much.

Because of the strong positive correlation between age and lifetime drug use for youths, age was included as an independent variable in the regression models, along with privacy. This made it possible to test the significance of the effect of privacy with age controlled. Lifetime prevalence of drug use was computed for each drug and for every possible combination of age and level of privacy. Each regression model was based on a maximum of 54 observations, one observation for each possible combination of age (12 to 17) and privacy (1 to 9). Each observation was weighted by the sample size for the age-privacy combination, to provide unbiased estimation. Some combinations did not occur in the samples. The dependent variable was the survey estimate of prevalence of lifetime use of the drug for the age and privacy combination. The model used can be described as follows:

$P = B_0 + B_1X_1 + B_2X_2 + e$, where

P = survey estimate of prevalence of a drug for the age-privacy combination

X_1= age (values range from 12 to 17)

X_2= privacy level (values range from 1 to 9).

The regression coefficients were computed for each of six drugs in both 1979 and 1982. Models were run using the procedure REG of the Statistical Analysis System (SAS Institute 1982). The statistical significance of privacy (after adjusting for age) was evaluated using partial F-tests ($H_0:B_2 = 0$, given B_0, B_1).

Once regression equations were obtained, expected values for the dependent variable were computed under each of two assumptions:

1) Assume every respondent had complete privacy (privacy = 1)

2) Assume every respondent had no privacy (privacy = 9)

The percent difference between these two expected values provides a standardized measure of the effect of privacy (vs. no privacy) on reported use of various drugs. The percent difference is computed by subtracting the estimated prevalence assuming no privacy from the estimated prevalence assuming complete privacy, and then by dividing by the estimated prevalence assuming complete privacy. It can be interpreted as an estimate of the proportion of lifetime users of the specific drug who would not report their use when interviewed with no privacy. The expected prevalence, assuming complete privacy, can also be compared with the actual prevalence obtained in the survey to estimate the degree to which drug use is underreported by youths in the National Survey as a result of a lack of complete privacy in every household.

RESULTS

Table 1 shows the proportion of youth interviews conducted at each privacy level in 1979 and 1982. It illustrates two conclusions.

First, interviewers in the survey have been successful in obtaining privacy. Well over half of all interviews had complete privacy, and over 80 percent of interviews were in the first three categories, which include "minor distractions." Second, the degree of privacy obtained in 1982 was very similar to that obtained in 1979. Thus, any reporting bias resulting from lack of privacy can be assumed to have little or no impact on trend analysis of National Survey data, assuming the effect of privacy is nearly constant.

TABLE 1. *Percent distributions of 1979 and 1982 National Household Survey on Drug Abuse sample of youths by privacy during interview*

	Percent Distribution	
Privacy Level	1979 (n=2,148)	1982 (n=1,538)
1 (Completely private)	60.8	57.9
2	8.6	9.6
3 (Minor distractions)	13.7	15.8
4	1.9	2.0
5 (Parent 1/3 time)	7.3	6.1
6	0.9	1.1
7 (Serious interruptions)	0.5	1.1
8	0.3	0.7
9 (Constant parent)	6.1	5.7

Table 2 shows the results of the regression analyses. Values for R^2 ranged from .35 to .86. These high R^2 values were primarily due to the age variable in the models. However, privacy was significant at the .1 level in 8 of the 12 regressions and at the .05 level in 4 of those 8. While the statistical significance of privacy is not overwhelming for the 1982 data, the consistency of the effect cannot be ignored. As table 3 shows, for 11 of the 12 regressions, privacy had a positive effect on reported drug use. The one case where the effect was negative was nonsignificant. Furthermore, the estimated effect of privacy is very large in some cases, indicating that if sufficient privacy is not achieved when conducting the interviews, drug use could be severely underreported. However, the high levels of privacy achieved in the National Survey (table 1) result in a minimal overall impact of this bias on that survey, as is indicated by the small differences between the first two columns of table 3.

TABLE 2. *Results of regression analysis of privacy and age with reported lifetime prevalence of drug use*

Year of Survey and Drug	R^2	Significance Level of Privacy in Model (H_0: $B_2 = 0$, Given B_0, B_1)
1979		
Cigarettes*	.60	.032
Alcohol	.73	.019
Marijuana	.86	.063
Cocaine	.55	.098
Hallucinogens	.66	.002
Psychotherapeutics*	.49	.061
1982		
Cigarettes*	.61	.002
Alcohol	.74	.293
Marijuana	.69	.208
Cocaine	.51	.700
Hallucinogens	.35	.627
Psychotherapeutics	.48	.084

*Self-administered form not used.

TABLE 3. *Estimates of lifetime prevalence of drug use for youths and the effect of privacy in interviews*

	Percent of Youths Ever Used			
Year of Survey and Drug	Reported in Survey	Assuming Complete Privacy	Assuming No Privacy	Percent Difference Between Complete Privacy and No Privacy (Privacy Effect)
1979				
Cigarettes*	54.2	56.0	44.9	19.8
Alcohol	70.3	72.0	61.4	14.7
Marijuana	30.8	32.0	24.5	23.4
Cocaine	5.3	5.9	2.3	61.0
Hallucinogens	7.0	7.9	2.2	72.2
Psychotherapeutics*	7.3	8.2	2.4	70.7
1982				
Cigarettes*	49.6	52.8	31.3	40.7
Alcohol	65.0	65.9	60.4	8.3
Marijuana	26.8	28.1	20.4	27.4
Cocaine	6.5	6.7	5.7	14.9
Hallucinogens	5.0	4.7	6.5	-38.3
Psychotherapeutics	10.3	11.3	5.1	54.9

*Self-administered form not used.

Although comparisons of the estimated effect of privacy between drugs and years may be outside the limits of this analysis, a few interesting differences are worth noting. The estimated effect of privacy was greater for cigarettes in 1982 than it was in 1979 (40.7 percent vs. 19.8 percent). The reverse was true for cocaine and hallucinogens, for which the privacy effect was greater in 1979. If these changes in the effect of privacy on reporting have really occurred, then trend analysis of National Survey data may be slightly biased. Adjustment for the bias results in a slightly smaller increase in cocaine prevalence between 1979 and 1982. The increase is statistically nonsignificant without this adjustment, and remains so after the adjustment. The decrease in cigarette prevalence, significant at the .05 level, becomes smaller after adjustment, and becomes nonsignificant. The nonsignificant decrease in hallucinogen prevalence becomes larger and significant at the .05 level after adjusting for the bias.

Cigarette questions in 1979 and 1982 and psychotherapeutic questions in 1979 were answered verbally by respondents, while all other drug questions (including psychotherapeutics in 1982) were answered using self-administered answer sheets. This should be kept in mind when comparing the estimated effect of privacy among the drugs and years, since underreporting would be expected to be greater without the self-administered form. While the estimated effect of privacy for psychotherapeutics was greater in 1979 using the verbal responses than it was in 1982 using the self-administered form, the difference is not large and could certainly result from random variation or error in the regression estimation. Of more interest is the overall increase in the percentage of youths using psychotherapeutics, which is quite likely to be due to the change in methodology, as suggested by studies showing that self-administered questionnaires produce higher reported prevalence of drug use than personal interviews (Hochstim 1967; Krohn et al. 1975).

DISCUSSION

The estimation procedure used here is admittedly imprecise. Nevertheless, the consistency of the direction of the privacy effect provides strong evidence that privacy is important in youth drug surveys, even when self-administered answer sheets are used. This result contradicts findings from an earlier study (Zanes and Matsoukas 1979), which concluded that the presence of parents in the room had no effect on reporting of drug use by youths. Different sample populations may be the cause of this difference. The National Survey is a nationally representative sample of youths ages 12 to 17 years, while the earlier study included only eleventh graders in a single high school located in a stable, mostly white middle- and upper-middle-class neighborhood. Also, the studies were conducted in different years.

One other issue must be mentioned regarding this study. Respondents were not randomly allocated to different levels of privacy. This raises the possibility that there is some other variable that is related both to privacy and to drug use.

This was investigated to the extent possible by analyzing the relationship between privacy and several variables, including sex, family size, race, geographic area, type of housing, and occupation of the head of the household, and then repeating the analysis using reported drug use with each of the variables. In general, most population groups that reported higher prevalence of drug use did not have significantly more privacy than lower prevalence groups. One exception was whites, who had slightly higher reported prevalence of cigarette and alcohol use than other races and also had more privacy during interviews. To investigate this relationship, the regressions were rerun on whites only. The privacy effect remained consistent, indicating that the privacy effect is independent of race. In some cases, population groups with higher reported prevalence of drug use actually had been interviewed with less privacy. For example, while males reported higher prevalence of drug use, they had been interviewed with less privacy than females. Thus, it appears that the relationship between privacy and reported drug use demonstrated by the regressions is not the result of the variables included in the survey.

In conclusion, it does appear that reporting of drug use by youths is affected by the degree of privacy during the interview, even when a self-administered answer sheet is used by the respondent. This underscores the importance of achieving maximum privacy when conducting drug surveys and raises questions regarding the validity of data from surveys in which adequate privacy was not obtained.

REFERENCES

Draper, N., and Smith, H. Applied Regression Analysis. New York: John Wiley and Sons, 1966. 407 pp.

Fishburne, P.M.; Abelson, H.I.; and Cisin, I. National Survey on Drug Abuse: Main Findings: 1979. Prepared for the National Institute on Drug Abuse by Response Analysis Corp. and the Social Research Group of the George Washington Univ. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1980. 169 pp.

Hochstim, J.R. A critical comparison of three strategies of collecting data from households. J Am Stat Assoc, 62:976-989, 1967.

Hubbard, R.C.; Eckerman, W.C.; and Rachal, J.V. Methods of validating self-reports of drug use: A critical review. Proc Am Stat Assoc:406-409, 1976.

Johnston, L.D. Techniques for reducing measurement error in surveys of drug use. In: Robins, L.N., ed. Studying Drug Use and Abuse. New Brunswick, N.J.: Rutgers University Press, in press.

King, F.W. Anonymous vs. identifiable questionnaires in drug usage surveys. Am Psychol, 25:982-985, 1970.

Krohn, M.; Waldo, G.P.; and Chiricos, T.G. Self-reported delinquency: A comparison of structured interviews and self-administered checklists. J Crim Law Criminol, 65:545-553, 1975.

Luetgert, M.J., and Armstrong, A.H. Methodological issues in drug usage surveys: Anonymity, recency and frequency. Int J Addict, 8:683-689, 1973.
Miller, J.D.; Cisin, I.H.; Gardner-Keaton, H.; Harrell, A.V.; Wirtz, P.W.; Abelson, H.I.; and Fishburne, P.M. National Survey on Drug Abuse: Main Findings: 1982. Prepared for the National Institute on Drug Abuse by the Social Research Group of The George Washington Univ. and Response Analysis Corp. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1983. 141pp.
Needle, R.; McCubbin, H.; Lorence, J.; and Hochhauser, M. Reliability and validity of adolescent self-reported drug use in a family-based study: A methodological report. Int J Addict, 18:901-912, 1983.
O'Malley, P.M.; Bachman, J.G.; and Johnston, L.D. Reliability and consistency in self-reports of drug use. Int J Addict, 18:805-824, 1983.
SAS Institute, Inc. SAS User's Guide: Statistics. 1982 Edition. Cary, N.C.: SAS Institute, Inc., 1982. 584 pp.
Single, E.; Kandel, D.; and Johnson, B.D. The reliability and validity of drug use responses in a large-scale longitudinal survey. J Drug Issues, 5:426-443, 1975.
Smart, R.G. Recent studies of the validity and reliability of self-reported drug use, 1970-1974. Can J Criminol, 17:326-333, 1975.
Smart, R.G., and Jarvis, G.K. Do self-report studies of drug use really give dependable results? Can J Criminol, 23:83-92, 1981.
Sudman, S., and Bradburn, N. Response Effects in Surveys: A review and synthesis. Chicago: Aldine, 1974.
Zanes, A., and Matsoukas, E. Different settings, different results? A comparison of school and home responses. Public Opin Q, 43:550-557, 1979.

AUTHOR

Joseph Gfroerer, B.A.
Division of Epidemiology and Statistical Analysis
National Institute on Drug Abuse
5600 Fishers Lane
Rockville, MD 20857

ISSUES OF VALIDITY AND POPULATION COVERAGE IN STUDENT SURVEYS OF DRUG USE

Lloyd D. Johnston
Patrick M. O'Malley

As with any survey technique, surveys of young people in school have both advantages and disadvantages for the quantification and study of drug abuse in the population.

ADVANTAGES OF STUDENT SURVEYS

In an earlier review of the use of survey methods in the drug field (United Nations 1980), Johnston listed the following advantages of surveys of special populations in institutions, including student surveys: First, there are considerable economies involved in being able to take clustered samples and in administering the instruments in groups. The lowered cost per respondent permits the collection of data from larger numbers of respondents, yielding more accurate estimates of prevalence and trends and a greater ability to conduct subgroup analyses. It also results in the collection of data on more users of the various substances, thus permitting the characterization and study of the users of more rare substances, such as cocaine or PCP, and even heavy users of certain substances, for example, daily marijuana users.

Insofar as drug use is particularly concentrated in the population of interest--which is certainly the case of people of secondary school and college age in contemporary North America--student surveys are capable of yielding a fairly high cost-benefit ratio by focusing on those more at risk and/or involved. Again, more users will be identified for study, given a particular sample size, than in a survey of the population at large. Finally, the population under study is already in identified institutions, giving them a particular accessibility for planned interventions.

Two other advantages, mentioned by Smart et al. (1980) in a review of student survey methodology, are that nonresponse rates tend to be low in school settings, and that the degree of anonymity obtainable is likely to be much greater than in other methods. The rate of nonresponse to the questionnaire among those present in the classroom is often under 1 percent (e.g., Johnston et al.

1982a), undoubtedly due in large part to the demand characteristics of the situation. This compares very favorably to the non-cooperation rates usually experienced in household or telephone surveys.

While household interviews sometimes involve the use of private answer sheets, which the interviewer promises not to read (Miller et al. 1983), obviously the interviewer knows the name and address of the person and, from the respondents' perspective, might abridge confidentiality by examining the answer sheets. (Similarly, in a telephone survey, the respondents' phone numbers are known, and for all they know, their names and addresses as well.) In a school survey, nearly complete anonymity can be convincingly given by the group collection of unidentified questionnaires. And if confidentiality, but not anonymity, is to be offered, this still can be done in a convincing way (Johnston 1980; Johnston et al. 1982a).

A related advantage, not mentioned in either of the two previous reviews, is that the young people are answering the sensitive questions about illicit behaviors without the proximity of their parents or other family members. While we know of no empirical research to date demonstrating a suppression effect of such proximity in household surveys, we would certainly hypothesize that there is one. A corollary hypothesis is that the effect would be greater if an adult member of the household is also being interviewed on the same topic--as happens in some household surveys (Miller et al. 1983)--since we think the young person would be concerned about being questioned later by that adult about what his or her answers were.

Another advantage that derives from the large sample sizes economically possible in student surveys is that the sample may be broken down into subsamples, each of which receives a partially different questionnaire. This permits the inclusion of many more variables in the study. Examples may be found in the Monitoring the Future survey and the more recent national student survey in Greece (Kokkevi 1984).

Finally, insofar as student surveys tend to make use of self-administered questionnaires, composed of precoded items, they also have the technical advantage over household interviews of using machine-readable forms, which considerably reduce the time and costs required to generate a computer-readable data file. This method can also reduce to near zero the error rate in data handling and processing, since it eliminates the human error involved in coding and keypunching (or otherwise entering) the data into a computer-readable form.

LIMITATIONS OF STUDENT SURVEYS

Having cited some of their virtues, let us hasten to add that they have their limitations, as well. Among the limitations of student surveys cited in those previous reviews (Johnston 1980; Smart et al. 1980) were the following: The population under study by definition excludes those out of the age range to be students, as well as those of equivalent age who already have left school. Those who tend to be absent from school more than average are another group that would be proportionally underrepresented--though not totally excluded from the sample--assuming there are no procedures to obtain absentees' participation after the time of the initial survey administration.

Other potential problems mentioned were the possible inadequacy of lists from which to draw a proper sample of schools from the larger universe being represented and the ever-present need to secure sufficient cooperation from schools to yield a reasonably representative sample of schools. In the United States and a number of other countries there are sufficient listings of schools from which to draw samples, but the autonomy of local school systems makes actually securing a representative sample a formidable task. (Of course, to the extent that representativeness is not a crucial feature of the research design--as is the case in some relational and evaluation design--this obstacle is far less serious, because researchers are free to "shop around" until they find enough schools willing to cooperate.)

Another limitation of school surveys, to the extent that they tend to use self-administered questionnaires, is that they cannot have as complex a branching of questions (that is, a branching in the sequence which is determined by the respondent's answer to prior questions). Some branching can be included in self-administered questionnaires, of course (Bachman et al. 1984), but too elaborate a sequence carries the danger that a respondent will become lost. However, it is worth noting that even in interviews complex branching sequences cannot be used in those segments to be answered by the respondent on private answer sheets, because the interviewer has to know the respondent's answers in order to do any branching.

THE "MONITORING THE FUTURE" STUDY

The remainder of this paper will be devoted to a more in-depth look at certain of these areas--in particular, the definition of the universe covered by student surveys, sampling issues raised in attempts to sample that universe accurately, the effects of omitting the same-age segment of the population not covered in the student universe (dropouts) as well as those absent on the day of the administration, and some issues having to do with the validity of the data gathered in student surveys. Many of the examples used will derive from the Monitoring the Future study, which is an ongoing series of annual surveys of high school seniors, with samples drawn to be representative of all seniors in public and private high schools in the coterminous United States in a given year (Bachman and Johnston 1978; Johnston et al. 1984b). One of the purposes of this series is to provide an accurate estimation of the prevalence and trends in the use of various substances (both licit and illicit) in this population. Estimates are also developed for high school graduates at later ages, up to age 28, using followup panels from each of the previously participating senior classes (O'Malley et al. 1984). However, because this cohort-sequential design feature is atypical of student surveys, we will not be discussing in this paper any of the issues it may raise.

DEFINITION OF THE UNIVERSE IN STUDENT SURVEYS

Student surveys usually have their universe defined in terms of students registered in particular kinds of schools at a particular grade level in a particular year; for example, seniors in academic high schools in 1984. Several different age levels are usually contained in varying proportions at any given grade, which makes comparisons with census and household data sometimes more difficult than might be supposed.

Grade Levels Encompassed

The grade level(s) chosen have methodological importance for several reasons, since grade level tends to be correlated with reading skills and the ability to follow directions, as well as with the proportion of the age group likely to be in school. In the developed countries, relatively large proportions remain enrolled even at the level of secondary school, though there are still substantial differences in school enrollment rates among countries. In the developing countries, the majority of a class cohort is likely not to be in school by the secondary level.

In Monitoring the Future we chose to focus on senior year, which corresponds approximately to ages 17 and 18, because: 1) it represents the end point in universal public education and the great majority of a class cohort are still in school; 2) it represents a "jumping off" point from which students make a great many different changes in environment and role status--ones which we want to study; 3) the students are old enough that the schools are less protective of them on sensitive subjects such as drug use, which increases our chances of securing a high enough school cooperation rate to develop a representative sample; 4) it represents an ideal point at which to take stock of the cumulative influences of family and school; and 5) it represents a good "check point" in the development process at which to measure drug involvement of various sorts. While there is good evidence to suggest that the incidence and prevalence of certain drugs continue to rise into the early twenties (O'Malley et al. 1984, Kandel, in press), the fact is that by the end of senior year nearly two-thirds of seniors in recent class cohorts have already had their initial experience with illicit drug use, and an even greater proportion have had their initial experiences with alcohol and cigarettes (Johnston et al. 1984b).

Types of Schools Encompassed

There tend to be a number of different types of schools in most countries that could be included in the sampling universe. Obviously, the more comprehensive the inclusion, the more generalizable are the findings. The Monitoring the Future study encompasses both private and public schools, as well as academic high schools and vocational high schools. A nationwide survey of drug use among secondary school students in Greece, currently underway, does the same (Kokkevi 1984). Seeking inclusiveness is particularly important if the student survey is to be one in a series intended to measure trends, since there is always the possibility that the distribution of students enrolled across the different types of schools may shift over time.

SAMPLING ISSUES

For the purpose of selecting samples that are representative of a given universe, a multistage sample is often used in school surveys. While the first stage might be the selection of schools, in a survey of a large geographical area, such as a country, there may be cost advantages to first selecting a set of primary sampling areas (PSA's) that usually correspond to counties, and which contain a population which is itself representative of the general population in the country. This is the procedure we use in Monitoring the Future--the reason being that the University of Michigan's Survey Research Center retains a permanent staff of interviewers who live in a fixed set of PSA's (which are used for various national surveys), thus making a field operation confined to those areas far less costly than one in which we would have to send our field teams to far-flung areas around the country. Little sampling accuracy is lost by using this first stage.

Selecting Schools

The second stage involves selecting schools within those areas, which requires as an initial step obtaining an enumeration of all schools within those areas from which to make a random selection. In the Monitoring the Future study we select the schools with probability proportionate to their size--that is, proportionate to the estimated number of seniors in the school. When schools are selected proportionate to size, the optimal sampling algorithm calls for equal-sized samples of students in each school in order for all students in the universe to have equal probability of selection. This method increases the number of large schools in the sample; that is important because otherwise only a few large schools in the sample would represent all students from large schools, and thus most students from large cities, since community size and school size are fairly highly correlated.

The alternative sample design is to give all schools equal probability of selection and then to take all students (or some fixed proportion of all students) in every school. The fact that a sizable portion of the sample would then come from just a few schools reduces sampling accuracy in general, even if school size were not confounded with degree of urbanicity.

In Monitoring the Future we draw schools with probability proportionate to size, but then, in those schools having more seniors than the number specified in the optimal sample design, we take most or all of those seniors anyway. We do this simply because it is often administratively more convenient for us, and for the schools, to take all seniors than to have to sample them, and the marginal cost of the extra cases tends to be very low since the interviewers are already in these schools. The extra cases increase the accuracy of the sample, but not as much as if they were being drawn from additional schools. If the school is very large (400 seniors), we will randomly sample classrooms, being sure that we have picked a set of classroom periods that contain all seniors and each senior only once. This is the third stage of sampling.

Corrective weighting must be introduced to correct for unequal probabilities of being chosen that are known to have been introduced at the different stages of sampling (Kish 1965). Assuming that proper randomization procedures are followed, and proper corrective weighting is introduced in the analyses of the data that eventually result, the resulting sample should be representative of the universe (leaving aside for the moment the nonresponse issue).

However, statistics based on the assumptions of a simple random sampling procedure cannot be applied to the results from such a sample. Corrections must be introduced to take into account the effects of clustering in the sample, that is, the effects of having drawn whole groups of students (in schools, and also within schools, in classrooms), rather than having drawn each student totally independently of all the others. Procedures exist for adjusting statistical tests for the loss of accuracy introduced by the clustered sample approach (Kish 1965).

Examples of the adjustments resulting in the Monitoring the Future study may be found in appendices on this subject in Bachman et al. 1984 and Johnston et al. 1981. In essence one can find a reduced "effective N"--that is, a sample size that, if based on a simple random sample, would yield sampling accuracy equivalent to the clustered sample actually obtained. The smaller the average cluster size, the less will be the proportional downward adjustment. In any case, the very large sample size in the Monitoring the Future study (N is approximately 17,000 cases per year) still yields a very respectable "effective N" for nearly all purposes.

THE EFFECTS OF SCHOOL NONPARTICIPATION

The study is designed in such a way that each year (after the first) the sample of schools consists of half participating for the first time, and half participating for the second time. In 1976 and subsequent years, participation rates for the new half-samples of schools have ranged from 66 percent to 80 percent. Half of the sample in each of these years consisted of repeat schools, ones that had participated in the previous year. The rates of repeat (i.e., second-year) participation range from 95 to 100 percent. Any schools that dropped out were replaced with substitute schools. These substitute schools were from the same geographic areas, from similar neighborhoods, and of similar size and racial composition. In the event of a refusal by the substitute school, a second (and if necessary, a third or fourth) substitute school was selected and invited to participate. Cooperation was obtained from an original or a substitute school in all but one or two instances each year. In the very few cases where no school was obtained, compensatory weighting of the data from similar participating schools was used to improve the population estimates.

It is reasonable to ask whether nonparticipation of some of the originally sampled schools is likely to have a significant effect on the findings. Insofar as population estimates of drug use and attitudes are concerned, the answer depends on two factors: the size of the refusal rate and the similarity of the substitute schools to the original schools they are replacing. With respect to the first factor, only between one-fifth and one-third (in early surveys) of the schools have been substitutes during any given year. With respect to the second factor, the substitutes are chosen to be as similar as possible to the original school. There is no particular reason to expect that the students in schools that refuse are greatly different from those in schools that agree to participate, since the reasons for school nonparticipation are based primarily on general policy issues and/or on somewhat happenstance events that are not likely to relate systematically to student drug use. In sum, the school refusal rate is not excessively high compared with other school-based studies, and the substitute schools seem likely to be quite similar to the refusal schools.

There is one additional point to be considered. Insofar as monitoring change is concerned, the effects of school nonparticipation should be minimal. Any systematic biases that might emerge (say, underrepresenting politically conservative districts) should be approximately replicated from year to year, so the trend data should still accurately reflect any major changes in drug use that might be occurring. A partial check on the adequacy of the sample schools can be made by comparing trend data based on the total sample with trend data based only on the half-sample that remains constant from one year to the next. Since this half-sample consists of the same set of schools, the trends cannot be affected by schools' participation or refusal. We have examined drug-use trend estimates, comparing the data from all schools with the data from only the matched half-samples. These estimates were extremely similar, suggesting that any error due to sampling of schools is constant.

THE EFFECTS OF OMITTING ABSENTEES

The proportion of students in the classroom at the time of administration who decline to participate generally tends to be extremely low (less than 1 percent) in Monitoring the Future. Thus, the only segment of the student universe omitted consists of those who are absent from school or class at the time of the administration. In Monitoring the Future, absentee rates range from 17 to 23 percent of the enrolled students (depending on the year), based on data from the teacher's class register from each classroom on the day of the survey.

Of course, it would be possible to try to collect data on subsequent occasions from the missing students, but we have judged this effort not worth the financial and administrative difficulty for both our interviewers and the schools. To be able to assess the effects on our estimates of drug use of omitting the absentees, we included a question in the study that asks students how many days of school they missed in the previous 4 weeks. Using this variable, we can place individuals into different strata as a function of how often they tend to be absent. For example, all students who been absent 50 percent of the time could form one stratum. Assuming that absence on the day of the administration is a fairly random event, we can use the respondents in this stratum to represent all students in the stratum, including the ones who happen to be absent that particular day. By giving them a double weight, they can be used to represent both themselves and the other 50 percent of their stratum that was absent that day. Those who say they were in school only one-third of the time would get a weight of three to represent themselves plus the two-thirds in their stratum who were not there, and so forth.

Table 1 shows the lifetime prevalence rates for the different drugs that result from the Monitoring the Future study, with and without this special weighting to correct for omitting the absentees. It also gives the prevalence rates deduced in this manner to exist among the absentees and the biases in the overall estimates that result from missing them. Tables 2 through 4 do the same things using annual, 30-day, and daily prevalence.

TABLE 1. *Lifetime drug use estimates for absentees and all high school seniors, 1981*

	Seniors Present	Absentees Only*	Seniors & Absentees*	Bias
Cigarettes	71.0	80.0	72.6	1.6
Alcohol	92.6	95.4	93.1	0.5
Marijuana	59.5	74.8	62.2	2.7
LSD	9.8	17.1	11.1	1.3
Other Psychedelics	9.1	17.0	10.5	1.4
Cocaine	16.5	27.8	18.5	2.0
Amphetamines	32.2	45.2	34.5	2.3
Quaaludes	10.6	17.9	11.9	1.3
Barbiturates	11.3	18.1	12.5	1.2
Tranquilizers	14.8	21.6	16.0	1.2
Heroin	1.1	2.2	1.3	0.2
Other Narcotics	10.1	16.3	11.2	1.1
Inhalants	12.3	17.9	13.3	1.0

*Estimates based on seniors' self-reported absences from school.

TABLE 2. *Annual drug use estimates for absentees and all high school seniors, 1981*

	Seniors Present	Absentees Only*	Seniors & Absentees*	Bias
Alcohol	87.0	91.5	87.8	0.8
Marijuana	46.1	62.5	49.0	2.9
LSD	6.5	11.6	7.4	0.9
Other Psychedelics	5.6	10.1	6.4	0.8
Cocaine	12.4	21.4	14.0	1.6
Amphetamines	26.0	38.4	28.2	2.2
Quaaludes	7.6	13.8	8.7	1.1
Barbiturates	6.6	12.2	7.6	1.0
Tranquilizers	8.0	13.1	8.9	0.9
Heroin	0.5	1.6	0.7	0.2
Other Narcotics	6.0	10.5	6.8	0.8
Inhalants	4.1	6.4	4.5	0.4

*Estimates based on seniors' self-reported absences from school.

TABLE 3. *30-day drug use estimates for absentees and all high school seniors, 1981*

	Seniors Present	Absentees Only*	Seniors & Absentees*	Bias
Cigarettes	29.4	41.3	31.5	2.1
Alcohol	70.7	80.3	72.4	1.7
Marijuana	31.6	46.9	34.3	2.7
LSD	2.5	4.8	2.9	0.4
Other Psychedelics	2.1	4.4	2.5	0.4
Cocaine	5.8	11.4	6.8	1.0
Amphetamines	15.8	24.8	17.4	1.6
Quaaludes	3.1	6.5	3.7	0.6
Barbiturates	2.7	5.5	3.2	0.5
Tranquilizers	2.7	6.1	3.3	0.6
Heroin	0.2	0.8	0.3	0.1
Other Narcotics	2.1	4.4	2.5	0.4
Inhalants	1.5	3.2	1.8	0.3

*Estimates based on seniors' self-reported absences from school.

TABLE 4. *Daily drug use estimates for absentees and all high school seniors, 1981*

	Seniors Present	Absentees Only*	Seniors & Absentees*	Bias
Cigarettes	20.3	31.0	22.2	1.9
1/2 Pack or More per Day	13.5	22.0	15.0	1.5
Alcohol	6.0	11.1	6.9	0.9
Marijuana	7.0	13.8	8.2	1.2

*Estimates based on seniors' self-reported absences from school.

It can be seen in these four tables that, while absentees as a group are deduced to have appreciably higher than average usage levels for all licit and illicit drugs, their omission does not depress any of the prevalence estimates in any of the tables by more than 2.7 percent because they represent such a small proportion of the total sample. Considering that a substantial portion of those who are absent are likely to be absent for reasons unrelated to drug use--such as illness and participation in extracurricular activities--it may be surprising to see the extent of the differences. In any case, from the point of view of instructing policy or public perceptions, the small "corrections" in tables 1 through 4 appear to be of little or no significance. (The correction across all 13 drugs in the lifetime prevalence table averaged only 1.4 percent.) Further, such corrections should have virtually no effect on cross-time trend estimates unless the rate of absenteeism were changing, and we find no evidence in our data that it is. Put another way, the presence of a fairly slight underestimate that is constant across time should not influence trend results. Should absentee rates change appreciably, then it could be argued more convincingly that such corrections should be presented routinely.

THE EFFECT OF OMITTING DROPOUTS

While school studies do not purport to represent those not in school, the concern is still raised about how accurate the estimates would be if taken to represent the entire class cohort, both those still in and those now out of school. Unfortunately, we cannot derive corrections from data gathered from seniors to impute the prevalence rates for dropouts, as we did for absentees, since we have no completely appropriate stratum from which to "sample." We do know from our own previous research (Johnston 1973), as well as the work of others (Kandel 1982), that dropouts have prevalence rates for all classes of drugs that are substantially higher than the in-school students. In fact, the dropouts may not be too dissimilar to the absentees. But again, because dropouts represent a fairly limited proportion of the age group, we would expect their omission to have relatively little effect on the overall estimates--even assuming they have substantially higher than average rates of use.

This is particularly true when one considers the range of reasons for becoming a dropout. In a report based on a recent NIDA technical review on the effect of omitting dropouts, Clayton and Voss (1982) note the range of reasons why young people leave school early, a number of which would not be expected to relate to drug use. Citing a followup study of 2,600 ninth-graders in California who were followed through their high school years, they note that of the 19 percent who were known to have dropped out, 2 percent left as the result of serious illness or accidents and 32 percent were categorized as educationally handicapped (Elliot and Voss 1974). Economic hardship and pregnancy also undoubtedly account for some additional proportion.

However, as with absentees, there remains little doubt that dropouts as a group have higher than average involvement in drugs. The question, then, is how to estimate the effect of their omission on the overall usage rate. To do this, two parameters must first be estimated: the proportion of the class cohort that is missing from school, and the estimated rate of use for the various drugs in that missing segment.

The proportion who fail to complete secondary school is about 15 percent based on Census data published for 1977 (U.S. Bureau of the Census 1978), which showed that the proportion of 20- to 24-year-olds who were not high school graduates was 15.4 percent. (Younger age brackets are more difficult to use because they include some who are still enrolled in high school.) Monitoring the Future probably covers some small proportion of the 15 percent, since it takes place a few months before graduation, and not everyone will graduate. On the other hand, perhaps 1 percent to 2 percent of the age group Census shows as having a diploma get it through a General Equivalency Degree and thus would not be covered in Monitoring the Future. (Elliot and Voss report this result for less than 2 percent of their sample.) So these two factors probably cancel each other out. Thus, we use 15 percent as our estimate of the proportion of a class cohort not covered.

Extrapolating To Dropouts From Absentees

To estimate the drug-use prevalence rates for this group we used two methods. One is based on extrapolations in which we assume that the difference between dropouts and the seniors who participated in the study is equivalent to 1) the difference between absentees and participating seniors, 2) one and one-half times that difference, and 3) twice that difference. The last we would consider a rather extreme assumption.

The second method involves using the best recent national data on drug use among dropouts--namely the National Household Surveys on Drug Abuse (Fishburne et al. 1980; Miller et al. 1983). While these surveys have rather small samples of dropouts in the relevant age range in any given year, they should at least provide unbiased estimates for dropouts still in the household population.

Using the first method of correction, tables 5 through 8 again give adjusted prevalence estimates for lifetime, annual, 30-day, and daily use respectively, this time correcting simultaneously for both absentees and dropouts. The three different assumptions about how different dropouts are from participating seniors are included in the table along with the resulting bias under each assumption.

Several things should be noted in these tables. First, under the assumption that dropouts are just like absentees, no prevalence rate is changed by more than 5 percent over the estimate based on seniors only, even with the simultaneous correction for both absentees and dropouts. The largest correction involves marijuana, with lifetime prevalence rising from just under 60 percent to 64 percent. The second point to note is that even under the most extreme assumption--which results in extremely high prevalence rates for dropouts on all drugs, for example 90 percent lifetime prevalence for marijuana--the overall correction in any of the prevalence figures for any drug remains less than 7.5 percent. Again, marijuana shows the biggest correction (7.5 percent in annual prevalence, raising it from 46 percent uncorrected to 54 percent corrected). As we would have expected, the biggest proportional change occurs for heroin, since it represents the most deviant end of the drug-using spectrum and thus would be most associated with truancy and dropping out.

TABLE 5. *Lifetime drug use estimates for dropouts and high school seniors, 1981*

	Seniors Present	Dropouts Only			Total Population (Seniors, Absentees, and Dropouts)			Bias		
		A*	B*	C*	A*	B*	C*	A*	B*	C*
Cigarettes	71.0	80.0	84.5	89.0	73.7	74.4	75.1	2.7	3.4	4.1
Alcohol	92.6	95.4	96.8	98.2	93.4	93.7	93.9	0.8	1.1	1.3
Marijuana	59.5	74.8	82.5	90.1	64.1	65.2	66.4	4.6	5.7	6.9
LSD	9.8	17.1	20.8	24.4	12.0	12.5	13.1	2.2	2.7	3.3
Other Psychedelics	9.1	17.0	21.0	24.9	11.5	12.1	12.7	2.4	3.0	3.6
Cocaine	16.5	27.8	33.5	39.1	19.9	20.7	21.6	3.4	4.2	5.1
Amphetamines	32.2	45.2	51.7	58.2	36.1	37.1	38.1	3.9	4.9	5.9
Quaaludes	10.6	17.9	21.6	25.2	12.8	13.3	13.9	2.2	2.7	3.3
Barbiturates	11.3	18.1	21.5	24.9	13.3	13.9	14.4	2.0	2.6	3.1
Tranquilizers	14.8	21.6	25.0	28.4	16.8	17.4	17.9	2.0	2.6	3.1
Heroin	1.1	2.2	2.8	3.3	1.4	1.5	1.6	0.3	0.4	0.5
Other Narcotics	10.1	16.3	19.4	22.5	12.0	12.4	12.9	1.9	2.3	2.8
Inhalants	12.3	17.9	20.7	23.5	14.0	14.4	14.8	1.7	2.1	2.5

*Estimates A, B, and C are derived by assuming that the difference between dropouts and seniors present, compared to the difference between absentees and seniors present, is:

A. the same as the difference for absentees,
B. half again as much as the difference for absentees,
C. twice as much as the difference for absentees.

TABLE 6. *Annual drug use estimates for dropouts and high school seniors, 1981*

	Seniors Present	Dropouts Only			Total Population (Seniors, Absentees, and Dropouts)			Bias		
		A*	B*	C*	A*	B*	C*	A*	B*	C*
Alcohol	87.0	91.5	93.8	96.0	88.4	88.7	89.0	1.4	1.7	2.0
Marijuana	46.1	62.5	70.7	78.9	51.0	52.3	53.5	4.9	6.2	7.4
LSD	6.5	11.6	14.2	16.7	8.0	8.4	8.8	1.5	1.9	2.3
Other Psychedelics	5.6	10.1	12.4	14.6	7.0	7.3	7.6	1.4	1.7	2.0
Cocaine	12.4	21.4	25.9	30.4	15.1	15.8	16.5	2.7	3.4	4.1
Amphetamines	26.0	38.4	44.6	50.8	29.7	30.7	31.6	3.7	4.7	5.6
Quaaludes	7.6	13.8	16.9	20.0	9.5	9.9	10.4	1.9	2.3	2.8
Barbiturates	6.6	12.2	15.0	17.8	8.3	8.7	9.1	1.7	2.1	2.5
Tranquilizers	8.0	13.1	15.7	18.2	9.5	9.9	10.3	1.5	1.9	2.3
Heroin	0.5	1.6	2.2	2.7	0.8	0.9	1.0	0.3	0.4	0.5
Other Narcotics	6.0	10.5	12.8	15.0	7.4	7.7	8.0	1.4	1.7	2.0
Inhalants	4.1	6.4	7.6	8.7	4.8	5.0	5.1	0.7	0.9	1.0

*Estimates A, B, and C are derived by assuming that the difference between dropouts and seniors present, compared to the difference between absentees and seniors present, is:

A. the same as the difference for absentees,
B. half again as much as the difference for absentees,
C. twice as much as the difference for absentees.

TABLE 7. *30-day drug use estimates for dropouts and high school seniors, 1981*

	Seniors Present	Dropouts Only			Total Population (Seniors, Absentees, and Dropouts)			Bias		
		A*	B*	C*	A*	B*	C*	A*	B*	C*
Cigarettes	29.4	41.3	47.3	53.2	33.0	33.9	34.8	3.6	4.5	5.4
Alcohol	70.7	80.3	85.1	89.9	73.6	74.3	75.0	2.9	3.6	4.3
Marijuana	31.6	46.9	54.6	62.2	36.2	37.2	38.5	4.6	5.6	6.9
LSD	2.5	4.8	6.0	7.1	3.2	3.3	3.5	0.7	0.8	1.0
Other Psychedelics	2.1	4.4	5.6	6.7	2.8	3.0	3.1	0.7	0.9	1.0
Cocaine	5.8	11.4	14.2	17.0	7.5	7.9	8.3	1.7	2.1	2.5
Amphetamines	15.8	24.8	29.3	33.8	18.5	19.2	19.9	2.7	3.4	4.1
Quaaludes	3.1	6.5	8.2	9.9	4.1	4.4	4.6	1.0	1.3	1.5
Barbiturates	2.7	5.5	6.9	8.3	3.5	3.8	4.0	0.8	1.1	1.3
Tranquilizers	2.7	6.1	7.8	9.5	3.7	4.0	4.2	1.0	1.3	1.5
Heroin	0.2	0.8	1.1	1.4	0.4	0.5	0.5	0.2	0.2	0.3
Other Narcotics	2.1	4.4	5.6	6.7	2.8	3.0	3.1	0.7	0.9	1.0
Inhalants	1.5	3.2	4.1	4.9	2.0	2.1	2.3	0.5	0.6	0.8

*Estimates A, B, and C are derived by assuming that the difference between dropouts and seniors present, compared to the difference between absentees and seniors present, is:

A. the same as the difference for absentees,
B. half again as much as the difference for absentees,
C. twice as much as the difference for absentees.

TABLE 8. *Daily drug use estimates for dropouts and high school seniors, 1981*

	Seniors Present	Dropouts Only			Total Population (Seniors, Absentees, and Dropouts)			Bias		
		A*	B*	C*	A*	B*	C*	A*	B*	C*
Cigarettes	20.3	31.0	36.4	41.7	23.5	24.3	25.1	3.2	4.0	4.8
1/2 pack or more per day	13.5	22.0	26.3	30.5	16.1	16.7	17.3	2.6	3.2	3.8
Alcohol	6.0	11.1	13.7	16.2	7.5	7.9	8.3	1.5	1.9	2.3
Marijuana	7.0	13.8	17.2	20.6	9.0	9.6	10.1	2.0	2.6	3.1

*Estimates A, B, and C are derived by assuming that the difference between dropouts and seniors present, compared to the difference between absentees and seniors present, is:

A. the same as the difference for absentees,
B. half again as much as the difference for absentees,
C. twice as much as the difference for absentees.

Extrapolating From Household Surveys

The second method of estimating drug use among dropouts was by using data from household surveys on dropouts versus those remaining in school. We conducted secondary analyses of the archived data from the 1977 and 1979 National Household Surveys. Analyses were restricted to the age range 17- to 19-years-old, since about 95 percent of the Monitoring the Future respondents fall in this range. Of course, the numbers of cases were small. In the 1977 survey there were only 46 dropouts and 175 enrolled seniors in this age group. In the 1979 survey 92 dropouts and 266 seniors were included.

Table 9 shows the differences observed between these dropouts and enrolled seniors for the lifetime prevalence and monthly prevalence of marijuana in both 1977 and 1979. Also presented in the same table for comparison purposes are the estimated differences between dropouts and seniors (including absentees) that were generated under each of the three different assumptions used in the method just discussed. As can be seen in table 9, the estimated differences from the household survey data come out at a level at or below the least extreme assumption made in the previous method (i.e., where dropouts are assumed to have the same drug use levels as absentees).

TABLE 9. *Differences between dropouts and high school seniors in prevalence of marijuana*

	Household Survey Differences			Seniors Survey, 1979 Estimated Differences		
Time	Seniors	Dropouts	Difference	A	B	C
Lifetime Prevalence						
1977	46.9	58.7	11.8	12.6	20.3	27.9
1979	53.8	66.6	12.8	12.6	20.3	27.9
Monthly Prevalence						
1977	31.0	30.4	-0.6	12.6	20.3	27.9
1979	30.9	41.8	10.9	12.6	20.3	27.9
Numbers of Cases						
1977	175	46				
1979	266	92				

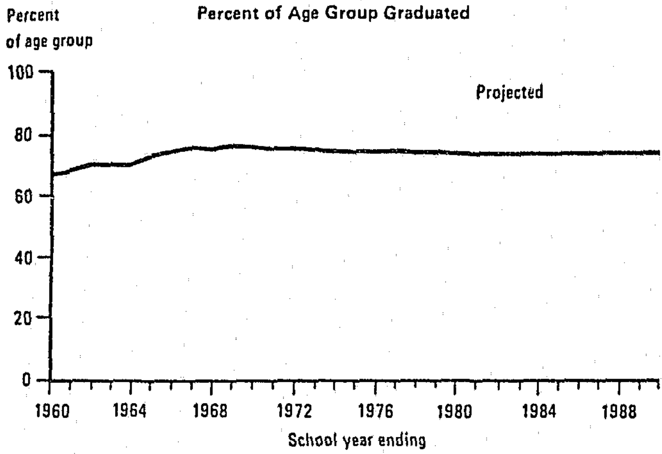
Note: The estimated differences for both lifetime and monthly prevalence in the seniors survey happen to coincide by chance.

While this may be comforting to the authors of the present paper, we must admit that we believe the household sample underrepresents the more drug-prone dropouts to some degree. Those without permanent residence and those in the prison population, to take perhaps the two most important examples, would be excluded from the sample coverage in a household survey. Thus, we are inclined to think that estimates closer to those made under assumption B--that dropouts are one and one-half times more likely to be users than absentees--may be closer to reality.

Again, we emphasize that there are a number of reasons for dropping out, many of which bear no relationship to drug use, including economic hardship in the family and certain learning disabilities and health problems. The extreme groups, such as those in jail or without a permanent place of residence, are undoubtedly a very small proportion of the total age group and probably even a small proportion of all dropouts. Thus, dropouts' prevalence rates would not change the prevalence estimates by much except in the case of the most rare events--in particular, heroin use. We do believe that, in the case of heroin use--particularly regular use--we are unlikely to get a very accurate estimate even with the corrections used in this paper. For the remaining drugs, we conclude that our estimates based on participating seniors, though somewhat low, are not a bad approximation for the age group as a whole.

Effects of Omitting Dropouts on Trend Estimates

Whether the omission of dropouts affects the estimates of trends in prevalence rates is another question, however. The relevant issues parallel those discussed earlier regarding the possible effects on trends of omitting the absentees. Most important is the question of whether the rate of dropping out has been changing in the country, since a substantial change would mean that seniors studied in different years would represent noncomparable segments of the whole cohort. Fortunately for the purposes of this study, the data published by the National Center for Educational Studies (NCES 1982) show that dropout rates stabilized in about 1968, following a period of slow decline, and have remained essentially stable up through 1980, which is the most recent year for which we have been able to locate published data. NCES projects the dropout rate to remain constant, as figure 1 illustrates. The reader should note that the statistic being traced in figure 1 is more an indicator of the dropout rate than a literal estimate, since it is based on how many 17- and 18-year-olds have completed school (excluding GED recipients) in each year. Clearly more of those two birth cohorts eventually will complete school as they become 18 (in the case of the 17-year-olds), 19, and so on.



Secondary schools graduated about 74 percent of the relevant age group, a proportion that has remained unchanged since the mid-1960s and is expected to be stable throughout the 1980s.

FIGURE 1. *Percent of 17- to 18-year-olds who have completed high school (excluding GED recipients)*

SOURCE: National Center for Educational Statistics. The Condition of Education, 1982 Edition, Chart 1.6, p. 19.

Given that there appears to be no sound evidence of a change in the dropout rate, the only reason that trend data from seniors would deviate from trends for the entire class cohort (including dropouts) would be if the constant proportion of those who have been dropping out for some reason showed trends contrary to those observed among seniors; even then, because of their small numbers, they would have to show dramatically different trends to be able to change the trend "story" very much.

There has been no convincing hypothesis offered for such a differential shift among dropouts. One hypothesis occasionally heard is that more youngsters are being expelled from school, or voluntarily leaving school, because of their drug use and that this explains the recent downturn in the use of many drugs being reported by the study (Johnston et al. 1984b). However, it is hard to reconcile this hypothesis with the virtually flat dropout rates over a 15-year period (through 1980), unless one posits a perfectly offsetting tendency for more completion among those who are less drug prone--hardly a very parsimonious set of explanations. Further, the reported prevalence of some drugs has remained remarkably stable throughout the life of the study, e.g., alcohol, opiates other than heroin, and the prevalence of some has risen, e.g., amphetamines, cocaine. These facts are not very consistent with the hypothesis of a recent increased rate of departure by the most drug prone. Certainly more youngsters leaving school in the 1980s have drug problems than was true in the 1960s. (So do more of those who stay in.) However, they still seem likely to be very much the same segment of the population, given the degree of association that exists between drug use and deviance and problem behaviors of various sorts.

In sum, while we believe there is some underestimation of the prevalance of drug use in the cohort at large as a result of the dropouts' being omitted from the universe of the study, we think the degree of underestimation is rather limited for all drugs (with the possible exception of heroin) and, more important, that trend estimates have been rather little affected. Short of having good trend data gathered directly from dropouts, we cannot close the case definitively. Nevertheless, we think the available evidence argues strongly against alternative hypotheses--a conclusion that was also reached by the members of the NIDA technical review on this subject held in 1982:

> ...the analyses provided in this report show that failure to include these two groups (absentees and dropouts) does not substantially affect the estimates of the incidence and prevalence of drug use (Clayton and Voss 1982 abstract).

VALIDITY ISSUES

Since the issue of validity is dealt with at some length in other chapters, we will address the subject only briefly here, and confine our comments to the Monitoring the Future study. By way of historical perspective, it should be said that we have come into this field with some skepticism about whether honest reporting could be secured through self-report on a topic as sensitive as illicit drug use. Some 15 years later we are firm believers that it can, at least in certain populations and with appropriate procedures. Certainly our extensive experience with the data from Monitoring the Future has convinced us of its fundamentally high quality.

To begin with, we have adopted a set of procedures that we think meets some necessary conditions for securing high cooperation and validity: namely, by convincing respondents that 1) there is a legitimate scientific or other reason for gathering the data, 2) they can answer in a situation that provides suitable privacy, 3) there are adequate procedures for continuing to protect confidentiality, and 4) those responsible for gathering and handling the data can be trusted. We will not go into detail regarding those procedures, since they have been described elsewhere

(Johnston et al. 1982a) and will be discussed at greater length in a forthcoming chapter (Johnston, in press) that reviews methods for increasing the validity of drug-use self-report data in general. Debriefing interviews in several high schools with some 100 students who had completed the questionnaire in the first year of the study convinced us that a high level of trust had been obtained.

But the most compelling evidence for us comes from the actual data generated. First, we have recently reported that, based on a series of three-wave panel analyses, the reliability and stability of our drug use measures tend to be quite high (O'Malley et al. 1983). Reliability estimates ranged in the eighties and nineties for cigarettes, alcohol, and marijuana lifetime prevalence measures, and in the seventies and eighties for illicit drugs other than marijuana taken as a group--certainly very respectable levels by most research standards.

In addition, because underreporting is the potential source of error of primary concern to us, we find it persuasive that in all classes surveyed so far, the majority of respondents have admitted using an illicit drug and fully two-thirds have made such an admission in the last couple of years, with 40 percent admitting to using an illicit drug other than marijuana. (In fact, the estimate rate for the absentee group exceeds 75 percent lifetime prevalence for illicit drug use based on respondents weighted according to their absenteeism (see table 1).) While troublesome for society, these exceptionally high proportions provide a kind of compelling evidence that, if there is systematic concealment, it is occurring only among a small minority (Johnston et al. 1982a, 1984b). The data from other questions on personal disapproval of drugs, and on the proportion of friends who use various drugs, give results that are highly consistent with the proportion self-reporting use. In fact, the aggregate level trends in friends' use and personal exposure to use--about which there is presumably much less motivation to lie--tend to parallel very closely the trends in self-reported use for the various drugs.

There is also strong evidence of construct validity to be found at the individual level of analysis in the relationships between use and a host of other variables that would be expected to relate in predictable ways to self-reported use. Among these are attitudes and beliefs about drugs, perceived availability of drugs, self-reported delinquency, truancy, religiosity, grades in school, evenings spent out of the house, etc. (Bachman et al. 1981; Johnston 1973). Most of these relationships are strong and replicable across graduating classes.

Finally, we have found that the missing data rates on the drug questions are only very slightly above normal for that point in the questionnaire where they occur, even though we have just instructed the respondents to skip the drug questions if they do not feel they can answer them honestly. For example, on the questions immediately preceding the drug section, the missing data rate averages about 2 percent, while for all of the drugs except marijuana the missing data rate runs from 2.5 percent to 3.0 percent (Johnston et al. 1982a). For marijuana, the missing data rate is slightly higher, at 4.5 percent, suggesting that marijuana use may be underestimated by about 2.5 percent. But overall, we find the absence of any substantial amount of skipping very reassuring.

CONCLUSIONS

The available evidence suggests to us that the noncoverage of absentees and dropouts has only modest implications for the estimation of overall prevalence rates and rather little implication for the estimation of trends in prevalence. Regarding the validity of self-report data in student surveys on drug use, we conclude that in the United States population, at least, it is possible with the proper procedures to secure data that have high reliability and show strong evidence of high validity as well. That is not to say that the data in such surveys are perfect, nor necessarily valid--since we think validity is highly dependent on the construction of the questions, the procedures under which they are administered, and the perceived intentions of the investigators. Further, we think that the validity of the trend data from such studies is dependent on the constancy of methods across time--particularly in question content, question context, field procedures, and timing of the survey during the year. But with the right procedures, and with the proper care given to keeping them constant across time, we believe such surveys can and do generate valid findings--on prevalence, trends, risk factors, and effects--which have considerable significance for the formulation and evaluation of policy, and for the advancement of scientific knowledge in this area.

REFERENCES

Bachman, J.G., and Johnston, L.D. The Monitoring the Future Project Design and Procedures. Ann Arbor: Institute for Social Research, 1978.

Bachman, J.G.; Johnston, L.D.; and O'Malley, P.M. Smoking, drinking, and drug use among American high school students: Correlates and trends, 1975-1979. Am J Public Health, 71(1):59-69, 1981.

Bachman, J.G.; Johnston, L.D.; and O'Malley, P.M. Monitoring the Future: Questionnaire Responses from the Nation's High School Seniors, 1982. Ann Arbor: Institute for Social Research, 1984.

Clayton, R.R., and Voss, H.L. Technical Review on Drug Abuse and Dropouts. National Institute on Drug Abuse. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1982.

Elliott, D., and Voss, H.L. Delinquency and Dropout. Lexington, MA: D.C. Heath-Lexington Books, 1974.

Fishburne, P.M.; Abelson, H.I.; and Cisin, I. National Survey on Drug Abuse: Main Findings 1979. National Institute on Drug Abuse. DHEW Pub. No. (ADM) 80-976. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1980.

Johnston, L.D. Drugs and American Youth. Ann Arbor: Institute for Social Research, 1973.

Johnston, L.D. Review of General Population Surveys of Drug Abuse. WHO Offset Pub. No. 52. Geneva: World Health Organization, 1980. 60 pp.

Johnston, L.D. The epidemiology of drug use and abuse: An overview of the research/policy questions and the methods appropriate for answering them. Paper distributed by Council of Europe's Pompidou Group. Document P-PG (82) 38. December 1982. 21 pp.

Johnston, L.D. Techniques for reducing measurement error in surveys of drug use. In: Robins, L.N., ed. Studying Drug Use. New Brunswick, NJ: Rutgers University Press, in press.

Johnston, L.D.; Bachman, J.G.; and O'Malley, P.M. Student Drug Use in America, 1975-1981. National Institute on Drug Abuse. DHEW Pub. No. (ADM) 82-1221. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1982(a).

Johnston, L.D.; Bachman, J.G.; and O'Malley, P.M. Student Drug Use, Attitudes and Beliefs: National Trends 1975-1982. National Institute on Drug Abuse. DHEW Pub. No. (ADM) 83-1260. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1982(b).

Johnston, L.D.; Bachman, J.G.; and O'Malley, P.M. Monitoring the Future: Questionnaire Responses from the Nation's High School Seniors, 1983. Ann Arbor: Institute for Social Research, 1984a.

Johnston, L.D.; O'Malley, P.M.; and Bachman, J.G. Highlights from Drugs and American High School Students, 1975-1983. National Institute on Drug Abuse. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1984(b).

Kandel, D.B.; Logan, J.A.; and Raveis, V.H. A comparison of former regular students, former school absentees, and school dropouts in a follow-up cohort of young adults. Unpublished manuscript, 1982.

Kandel, D.B. Developmental patterns of the use of legal, illegal, and medically prescribed psychotherapeutic drugs from adolescence to young adulthood. In: Jones, C., and Battjes, R., eds. The Etiology of Drug Abuse: Implications for Prevention. National Institute on Drug Abuse. Research Monograph, in press.

Kish, L. Survey Sampling. New York: John Wiley & Sons, 1965.

Kokkevi, A. Personal communication, 1984.

Miller, J.D., et al. National Survey on Drug Abuse: Main Findings 1982. National Institute on Drug Abuse. DHEW Pub. No. (ADM) 83-1263. Washington, D.C.: U.S. Govt. Print. Off., 1983.

National Center for Educational Statistics. The Condition of Education: 1982 Edition. DHEW Pub. No. NCES-82-400. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1982.

O'Malley, P.M.; Bachman, J.G.; and Johnston, L.D. Reliability and consistency in self-reports of drug use. Int J Addict, 19:805-824, 1983.

O'Malley, P.M.; Bachman, J.G.; and Johnston, L.D. Period, age, and cohort effects on substance use among American youth. Am J Public Health, 74:682-688, 1984.

Smart, R.G.; Hughes, P.H.; Johnston, L.D.; et al. Drug Use Among Non-Student Youth. Geneva: World Health Organization, 1980.

United Nations Division on Narcotic Drugs. Manual on Drug Abuse Assessment: Part II, the Use of Population Surveys. 1980.

U.S. Bureau of the Census. School Enrollment--Social and Economic Characteristics of Students, October 1976. Current Population Report Series P. 20, No. 319. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1978.

AUTHORS

Lloyd D. Johnston, Ph.D.
Patrick M. O'Malley, Ph.D.

Institute for Social Research
University of Michigan
Ann Arbor, MI 48106

SAMPLING AND COVERAGE DIFFICULTIES IN CANADIAN DRUG USE SURVEYS AND EFFORTS TO AVOID THEM

Reginald G. Smart

In Canada, surveys of illicit drug use and abuse have been conducted since the late 1960s. The first surveys were of school populations; not until about 1973 were any general population surveys done. Since then, much experience has been gained with a variety of assessments and survey techniques for drug abuse. We continue to have problems with sampling and with coverage of the whole population of students and adults. We have made some efforts to develop methods of dealing with some of the problems. Total coverage, however, remains elusive.

SCHOOL SURVEYS IN ONTARIO: SAMPLING AND COVERAGE PROBLEMS

The first drug use surveys in Ontario were studies of the Toronto school population (Smart and Jackson 1969). These were done every 2 years from 1968 to 1974. In 1977, it was decided to expand the sample to include all school districts in Ontario (Smart et al. 1983). This made it much more difficult to do the sampling and to collect the data, since there were 200 rather than only 5 school boards involved, and they extended over a wide geographic area. Currently, data are collected every 2 years, with the last survey in 1983.

Ontario is a very large province; it measures 700 miles from east to west and 1,200 miles from north to south. The northern parts have severe winters. Unfortunately, we determined early in the surveys that the best month for surveys is February. It avoids all holidays, school events, and examinations but is the worst month for travel. Originally we had the survey data collected by people travelling from Toronto. We now use a survey research center that has a local field staff, so there are fewer problems with survey staff not arriving. Snowstorms still reduce the "at school" population on the day of the survey, so we try to reschedule the survey if the weather is bad. Also, we do not include small, remote boards of education in the sampling frame. Those northerners left out constitute only 7 percent of the target student population.

Our coverage is incomplete in other ways. The school system in Ontario is very complex. There are both Ontario Public and Separate (Roman Catholic) School Boards in most areas. The separate boards include only students in grades 1 to 10. Those in grades 11 to 13 are "private school" students. We include public and separate schools up to grade 10, but only public schools in 11 to 13. This probably increases our reported illicit drug use rates since students in Catholic schools have somewhat lower rates of use. We also leave out all other private and church-related schools, but they educate a very small segment of the young population. It is not possible to say how this affects results.

A variety of types of schools and classes have been excluded from the provincial survey because of expected problems, such as special education classes, those students institutionalized for health or correctional reasons, and schools on Indian reservations. Schools on Canadian Forces (Army and Airforce) establishments also are excluded, but there are only a few of these. Probably, these exclusions tend to leave out heavier-drinking and drug-using populations. There are indications that those in our prisons and youth correctional institutions do have higher rates of drug abuse. The same is true of Indians living on reservations. However, about half of the Indians in Ontario do not live on reservations, and they would be included in the sample. Probably, our exclusions from total coverage serve to decrease the reported rates of drug and alcohol use.

Our sample is based on a stratified single-stage cluster probability sample design. The data are weighted to take into account variable sampling fractions and nonresponse by selected classes and students. The target population is students in regular programs in public and separate school boards in grades 5,7,9,11, and 13. We require a sample of 8,000 in order to get a minimum sample of 5,000 students. The sample is stratified into 4 regions and 5 grades, resulting in 20 strata or area/grade clusters. We used projected enrollment figures for 1983 based on those for 1980 to 1983 because we could not get the up-to-date figures early enough. Usually the projections are good. A probability sample is independently selected from within each of the 20 strata.

The sampling units are "homerooms" either as they actually exist in schools or on the basis of average size in the relevant stratum. In anticipation of refusals to participate we selected additional homerooms in each stratum.

When we started our provincial survey, we estimated, from Toronto data, class sizes as 25 for grade 13 and 30 for other grades. These estimates were too high, especially in the north, and probably resulted in some undersampling there (this area has many drinking problems). We now base the estimate of homeroom size on what we found last time in each stratum; there is less guesswork and it works out more accurately. In a few schools there are no homerooms, in that students have each class with a different set of students. We constructed some so that students could complete the questionnaire in groups of the right size--usually 25 to 30 students.

Our study depends greatly on cooperation by school boards and principals. In Ontario the principal has absolute power over such things as external surveys, even if the board approves them. Sometimes we have trouble getting the cooperation of school boards. It seems that the level of cooperation is falling. In our 1981 survey, 91 percent of boards approached agreed to the survey, but only 82 percent agreed in 1983. There is an increasing tendency for boards to say there are too many surveys and that drug use surveys tend to promote drug use. Others complain about the specifics of the questionnaire (too long, too difficult). In our 1983 study there were 10 boards that refused and 31 that cooperated. We replace boards that refuse but concern is growing that perhaps the cooperative boards are interested in the drug problem because they have an abnormally large number of student users. Sometimes boards will approve the survey overall but refuse the participation of grade 5 or 7 students on the basis that they are too young to be using any drugs and that use should not be encouraged.

In the past we did not replace boards. In the 1981 survey, a board in Western Ontario refused too late to allow replacement, and this resulted in loss of an entire stratum. Since then we have always allowed board replacement and we ask more boards than we will need so refusals are less important.

In Ontario, principals or headmasters can refuse to allow their school or certain classes to participate. In our early surveys, we allowed principals to designate classes to participate, but for some time we have insisted that we select at random. If principals do not agree, we drop the school and replace it. In fact, problems with individual schools are decreasing. In 1981, 75 percent of selected classes participated in the survey, but in 1983 some 97.8 percent did. We think the improvement occurred because we do more followup work with schools. We send schools copies of the survey report and offer to make community development help available to improve drug education programs. In the past, schools refused because they saw no benefit to the school from the survey. Others complained that there were too many surveys. We have tried to be helpful to schools by giving them advice about drug education programs. Also, we have explained that our survey typically requires few students per board from each grade level (one to five classes in all). In contrast, most surveys require large numbers of students.

A further problem with coverage of our target sample involves parental consent forms. Almost all of our students are below the age of 18 and hence, legal minors. By law, schools stand "in loco parentis" but some schools are not willing to mandate a drug use survey for younger students. Older students do not need parental consent. It is difficult to predict whether boards will require consent forms; if they do the forms are sent home with students and are to be returned in a few days. They add to the time and expense of the study; students without forms cannot participate. An earlier comparison indicated that schools with a low rate of forms returned had lower rates of reported illicit drug use. We make special efforts such as frequent reminders and detailed explanations of the study to get the forms returned. About 95 percent of them are approved and returned. Fewer are returned in metropolitan Toronto and for students in grades 5 and 7. The low rate of return in Toronto probably reflects the large population of parents who do not read English or French very well. Younger students probably lose the forms or parents think drug use surveys are inappropriate for them. Parents are more likely to refuse participation if the student is using drugs; there may be some concern that their children will be identified by school authorities.

Some of these results are similar to those of Kearney et al. (1983), who found that requiring parental consent forms caused a large reduction in sample size with an overrepresentation of whites and an underrepresentation of blacks and Asian Americans. Kearney et al. (1983) reported a sample reduction of 50 percent in their Seattle and Portland studies. However, this must reflect very inadequate followup by school officials since we find no more than a 5 percent reduction because of consent forms. Almost all parents in our studies who are asked for their consent give it. We send our consent forms home with students and they return them, whereas in Kearney et al.'s study, the consent forms were mailed.

A last problem concerns students who are not at school on the day of the survey. About 7 to 10 percent of students are away on any given day. The rate is higher in bad weather, especially during snowstorms and near holidays. Students may be away because of illness, family responsibilities, part-time work, or a dislike of school. Dropouts usually develop a pattern of infrequent attendance before actually leaving school completely. School absences are higher among males, those with lower grades, and those in the higher grade levels. Much research (e.g., Haberman et al. 1972) shows that students not at school on the survey day have relatively high rates of drug abuse. We have not done any studies of those not at school for the drug survey, because absenteeism rates are not very high and the results of such studies are predictable.

We have studied early school dropouts. In 1978, we studied 292 young people ages 14 to 18 during a household survey of adults. The young people were inhabitants of the same house where an adult was interviewed. We left a questionnaire to be filled in and returned later. Almost all were returned on time and complete. Table 1 shows some of the results. In the household survey students more often reported somewhat more cannabis use than did nonstudents (36.6 percent compared to 32.9 percent). However, both reported far more cannabis use than did students in a school survey a few months before. Results were different for tranquilizers; students in the school survey reported more use. These results were different from those in Mexico City and Chandigarh, where nonstudents reported more drug use (Smart et al. 1981). It appears that students identified in a household survey report more cannabis use than those in a school survey. Probably we got some students in the household survey who would have been absent on a survey day.

TABLE 1. *Comparison of school and nonschool populations in reported drug use (percent using in past 12 months)*

	Percent Users		
	Household Survey		School Survey
Drug	Students (N=257)	Nonstudents (N=173)	Students (N=4,794)
Cannabis	36.6	32.9	25.1
Tranquilizers	2.7	1.7	8.9

In summary, most but not all sources of noncoverage in our school survey probably lead to an underreporting of the true use of drugs such as cannabis. Probably our results are conservative underreports and we miss a number of heavier drug users. Our surveys do give a very good sampling of the school population by grade and geographic area (within 4 percent of expectancy), based on known demographic characteristics of the school population. Over 85 percent of students selected participated in 1983. If any students are missed, they are more likely to be drug users. Of course we would like to be completely accurate, but it is far better to be able to describe our drug use data as "underreports" than "overreports" or exaggerations. If we exaggerate the extent of drug use it is likely to be seen as self-serving and alarmist. Hence, the message would more likely be discounted. Frequently it is very difficult to know what problems a particular survey has had in sampling or other coverage problems. Details of such matters are often left out of published reports, but no real assessment of the reported drug use rates can be made without them.

Some types of noncoverage can be reduced by increasing staff time and commitment, e.g., to get consent forms returned. Other kinds can be reduced by increasing the survey's value to schools and school boards. We are concerned that in the long run refusals by boards will make studies difficult; hence we now offer consultation from Addiction Research Foundation community consultants. We also make the data for each board available to the trustees, but each board gets only its own data.

SAMPLING PROBLEMS IN HOUSEHOLD SURVEYS

Similar problems to those in school surveys arise with surveys of general populations. Most general population surveys--our own included--are of people living in houses and apartments. In some cases, even apartments cannot be included because of management's security concerns. The surveys also leave out people in prisons, hospitals, old age homes, hostels for the mentally ill, halfway houses for alcoholics, therapeutic communities for drug addicts, and many other institutions. Transients, street people, and skid row habitues are also missed, as are students in residence and people living on military bases.

These exclusions almost certainly reduce our reported drug use figures from their true value. Among those approached, there are problems in gaining completed interviews. Noncooperation rates in our unpaid surveys are 20 to 25 percent. We get the least cooperation from wealthy, upper-class residents who wish to protect their privacy and from the foreign-born, who often worry about surveys and government snooping into their affairs. We can improve the rates for foreign-born respondents by having interviewers who can speak different languages, but covering all of the possibilities is very difficult.

Many of our drug-use surveys have produced interview samples that are somewhat overweighted with older people. Older persons tend to be more often at home and more tolerant of long interviews, as they seem to have time on their hands. We have also found it difficult to interview sufficient numbers of young males without a large number of call-backs. Most of our interviewing is done in the evenings when young men are often not at home. We know that young males are more likely than others to be heavy drinkers and users of illicit drugs. Also, we expect that many of those who are out are heavier drinkers and drug users. Perhaps they are at bars, pubs, or parties where drugs are being used. In our survey (Smart and Goodstadt 1978) we have oversampled young males (aged 18 to 20) and have done the interviewing on Saturday morning. This is the best time to find young males at home, especially in the colder times of the year.

The results are shown in table 2. We cannot make an exact comparison between the expanded young male sample and the adult sample. However, it is obvious that the young males are a group containing many users of marijuana (41 percent) and stimulants (6 percent) and that virtually all are drinkers. If this group is missed in surveys, we will have underestimated drug use figures and lost contact with a special target population for drug education efforts.

TABLE 2. *Percentage of Ontario adult and expanded male samples using various drugs in past 12 months (1977 data)*

	Percent Users by Type of Drugs				
Age	Sleeping Pills	Stimulants	Tranquilizers	Marijuana	Alcohol
18 to 20 (males)	3.0	6.0	8.0	41.0	93.0
18 to 29	4.8	2.9	8.3	23.7	87.5
30 to 49	5.0	0.9	13.9	3.5	88.6
50+	16.4	1.3	17.3	0.7	69.8

Poor coverage in drug and alcohol surveys has been attributed to sampling problems and nonresponse. However, selective reporting and forgetfulness by respondents are also difficulties that cannot be easily overcome. We know that heavy drinkers are very likely to underreport their consumption in household surveys. Such surveys do not account for more than 40 to 60 percent of the known consumption of alcoholic beverages when sales figures are used.

We have experimented with an informant method (Smart and Liban 1982) to try to improve the estimates from surveys. In the informant method, selected individuals report the drinking practices (and amounts) for groups known to them. They do not report on their own drinking as in the usual surveys. This avoids the problem of respondents being asked to admit to overdrinking about which they probably feel guilty.

The participants or informants are selected to reflect the major occupational and geographic strata in a society. For example, in our study there were 30 groups including managerial/professional categories, secretarial/sales, industrial workers, housewives, students, farmers, and retired persons, with the number of groups of each in proportion to their representation in the population. The groups met once for about 2-1/2 hours to discuss answers to a 56-item questionnaire on drinking practices and attitudes. Each group reported only on drinking in the occupation stratum that they represented. Results are then aggregated for the 30 groups to provide a picture of drinking in the society as a whole.

The results indicated that, compared with a standard household survey done in the same area, the informant method gave better results. It reported higher rates of drinking and heavy drinking and gave per capita consumption figures close to those from alcohol sales (8.26 liters of absolute alcohol per year compared to 10.23 in sales figures). The survey method, as expected, gave poor estimates of per capita consumption (only 3.94 liters).

Because the method is cheap and more accurate than survey methods, it could be used in a variety of ways. It could be used to study heavy-using populations and to study drinking in developing countries without the resources for surveys. We have not applied it to drug-use surveys, but it should be useful in school studies where there is concern about noncoverage. Results from the informant method and a standard school survey could be compared. Unfortunately there are no good sales figures, as there are with alcohol, with which to validate the method.

In summary, coverage and sampling problems still plague school and general population surveys. Most of the problems work to decrease the estimates of illicit drug use and heavy drinking. We need continued research on new methods of reducing coverage problems. We also need more reporting of the details of sampling plans and how well they were actually achieved in practice.

REFERENCES

Haberman, P.W.; Josephson, E.; Zanes, A.; and Elinson, J. High school drug behavior: A methodological report on pilot studies. In: Einstein, S., and Allan, S., eds. Proceedings of the First International Conference on Student Drug Surveys. New York: Baywood Publishing Co., 1972. 346 pp.

Kearney, K.A.; Hopkins, R.H.; Mauss, A.L.; and Weisheit, R.A. Sample bias resulting from a requirement for written parental consent. Public Opin Q, 47:96-102, 1983.

Smart, R.G., and Jackson, D. A Preliminary Report on the Attitude and Behavior of Toronto Students in Relation to Drugs. Toronto: Addiction Research Foundation, 1969. 71 pp.

Smart, R.G., and Goodstadt, M.S. Alcohol and Drug Use Among Ontario Adults: Report of a Household Survey, 1977. Substudy No. 957. Toronto: Addiction Research Foundation, 1978. 11 pp.

Smart, R.G., and Liban, C.B. Alcohol consumption as estimated by the informant method, a household survey and sales data. J Stud Alcohol, 43:1020-1027, 1982.

Smart, R.G.; Mora, M.E.; Terroba, G.; and Varma, V.K. Drug use among non-students in three countries. Drug Alcohol Depend, 7:125-132, 1981.

Smart, R.G.; Goodstadt, M.S.; Adlaf, E.M.; Sheppard, M.A.; and Chan, G.C. Preliminary Report of Alcohol and Other Drug Use Among Ontario Students in 1983, and Trends Since 1977. Toronto: Addiction Research Foundation, 1983. 23 pp.

AUTHOR

Reginald G. Smart, Ph.D.
Addiction Research Foundation
33 Russell Street
Toronto, Ontario
Canada

DYNAMIC SIMULATION MODELS: HOW VALID ARE THEY?

Raymond C. Shreckengost

VALIDITY, PURPOSE, AND CONFIDENCE

In a strict sense, the subject of dynamic simulation model validity can be treated throughly and quickly: there are no fully valid models because all models are something less than the object, or system, being modeled. For example, millions of people have a conceptual model of the President, but, like fingerprints, no two of these models is exactly the same. Further, none matches precisely every detail of the real system. The same reasoning applies to all types and kinds of models of drug abuse.

In a practical sense, we are concerned with usefulness rather than validity. Does the model serve the purpose for which it was intended? Is it helpful? Thus, the developer's or user's purposes must be kept in mind in evaluating a model's usefulness, or validity. Criticisms of models also should reflect this perspective.

Much depends on the purpose for which the model is developed--for example, the choice of the level of detail used in the model. Just as micrometers are not used to measure intercity distances to a tenth of a mile, explicit modeling of each household would be equally absurd in a model treating the gross behavior of drug abuse systems. The selection of an appropriate level of detail, problem boundaries, and similar considerations constitute the "art" aspect of dynamic simulation model development.

Validity, or usefulness, lies in the subjective view of the user. We think of models as valid when they can be used with confidence. So, this paper focuses on how we can gain confidence in dynamic simulation models. In particular, it considers confidence or validity tests as they relate to a particular dynamic simulation method, System Dynamics (Forrester 1961, 1975; Roberts 1978).

These tests, however, are equally valid for other simulation techniques. They have evolved from nearly 30 years' experience of the inventor of System Dynamics, Jay W. Forrester (Bell and Senge 1980; Forrester and Senge 1980).

MODEL STRUCTURE TESTS

Because the foundation for model behavior is the model's structure, the first test in validating a model is whether the structure of the model matches the structure of the system being modeled. Every element of the model should have a real-world counterpart, and every important factor in the real system should be reflected in the model. Although this may seem like a simple, obvious test, it may not be so. For example, descriptions of how all of the structural parts of real systems are tied together rarely exist. More often than not, such descriptions must be based on the concepts, or mental models, of people familiar with the system. Further, important parts of some systems may lie unrecognized prior to modeling. During the development of a model dealing with the effects of heroin imports into the United States, for example, the key factor in the system, the relative abundance of heroin, was not immediately identified (Gardiner and Shreckengost this volume). Thus, the art of model building may, at times, entail discovery and invention.

This approach differs strongly from "Let's collect lots of data and then see what they tell us." Structure, like many other System Dynamics model elements, exploits judgment, experience, and intuition. Data play a secondary role.

Model Parameter Tests

The model's parameter values are a specific area for testing. Parameter values in a model often may be tested in a straightforward manner, e.g., against historical data. However, in dynamic simulation models of social systems the desired data may be unavailable, in an inappropriate form, or incorrect. There may be elements that are not usually quantified, but that are critical to the system being modeled. These elements must be included in the model. If prejudice, for example, is an important element, it must be included in the model, and its relationship to other pertinent parts of the system must be specified quantitatively. Many required parameter values may not exist and must be developed. In the heroin model, data and descriptions relating to heroin's relative abundance were initially absent. On the other hand, some available, apparently reasonable and acceptable, data on heroin imports turned out to be unreasonable and unacceptable when employed in the model. The point is that dynamic simulation model parameter values, from whatever source they may be derived, are subject to a rigorous and demanding environment. These values contribute significantly to confidence in the model when the specified parameter values are reasonable and consistent with whatever supporting data might exist.

Boundary Adequacy Test

If a model is focused on the heroin system in New York City, it will not generate national behavior. Conversely, a national heroin model is not likely to replicate the behavior of local systems. Model boundaries must match the purpose for which the model is designed, if the model is to be used with confidence: that is, the model must include all of the important factors affecting the behavior of interest. In practice, boundaries tend to shift as the developers' and users' understanding of a problem evolves with the model's development. As model purpose shifts, changes in the model's boundaries may be required.

In many problems, a simple model with limited boundaries may be expanded, or disaggregated, from time to time, as the model is used to address problems in greater detail. When this occurs, careful attention must be given to indirect effects, which may not be obvious. Suppose, for example, a model treating United States heroin users as a homogeneous group is disaggregated to identify users of small, medium, and large amounts. This will change the user boundaries, and associated changes will be needed in the consumption boundaries.

If the model boundaries are improper, or inadequate, the model's validity is degraded. However, criticism of dynamic simulation models aimed at boundary issues frequently reflects different notions about the model's intended use or purpose. For example, criticism of the user boundary in a model treating users as a homogenous group may ignore the fact that the grouping is consistent with the purpose of the model. But, as explained above, if the purpose is to account for different classes of users, a boundary change is required to account for the change in purpose.

Extreme Conditions Test

A less obvious test relating to model structure involves the effects of extreme conditions. The ability of a model to function properly under extreme conditions contributes to its utility as a policy evaluation tool as well as user confidence. Testing to extreme conditions may easily be overlooked or brushed aside in the hectic environment of early model development. Subsequently, this oversight may degrade model performance: subtly under normal conditions and significantly when the model is used to answer "What if?" questions that fall outside the operating regions emphasized in early development.

Again, the heroin import model provides a good example. In the past, heroin imports have been, roughly, 5 metric tons per year. They have not fallen to zero, nor have they soared to 10 or 20 tons. Consequently, during the model's development parameter values covered the range of import variations that were of immediate interest, say, 3 to 7 tons. If these initial values only were retained in later versions of the model, the model would show a residual, sizable user population even if imports were reduced to zero. At the other extreme, the number of users would reach an understated upper limit in the presence of a very large heroin supply surplus. The point is that model validity is enhanced if the region within which the model was originally designed to operate is extended so the model generates plausible behavior conditions outside the initial region. For example, the user population should be zero when imports are zero.

Tests under extreme conditions may also expose structural faults or inadequacies and incomplete or erroneous parameter values.

MODEL BEHAVIOR TESTS

Behavior Replication Test

The tests relating to model behavior are less technical and, for many users, more appealing and convincing than the structural tests. Foremost among these tests is the comparison of model behavior with the behavior of the system being modeled. A model whose behavior has little, or nothing, in common with that of the system of interest generates little, or no, confidence.

Where historical time series data are available, the model must be capable of producing similar data. That is, if the model's initial conditions are matched to the state of the system being modeled at some time in the past, the model's behavior should parallel the historical data from that time to the present. In this test, it is again important to keep in mind the purpose of the model--including the time span of the areas of behavior that are of interest. Further, judgment must be exercised about how closely the model's behavior should match the historical data, since historical data are less than perfect, and, sometimes, far from perfect. It is not at all uncommon for models to illuminate erroneous data. Where historical data are very poor or nonexistent, the test may be one of reasonableness.

Given the imports over 10 years or so, the heroin model (Gardiner and Shreckengost,this volume) generates heroin purity and price values that match well with the historical data for these parameters. Further, it also produces heroin-related death figures that match the historical data closely. The closeness of the model-historical correspondence is quite surprising, given the difficulties inherent in collecting and processing the data that the historical time series represents.

Purity, price, and deaths can be defined and measured with relative ease compared to the heroin user population. The model generates user population values against a strict, limited definition of a heroin user. Here, no parallel historical data exist, and the test becomes one of reasonableness considering the purpose for which the model was developed. Subsequent to its initial development, this sector of the model has been detailed to accommodate users with different consumption habits and varying responses to the abundance of heroin. Although this is intuitively more satisfying, there is still no opportunity for a confirming historical test.

Anomalous Behavior Test

When model behavior does not replicate the behavior of the real system, model structure, parameter values, boundaries, or similar factors are suspect. Something may have been omitted, improperly specified, or assigned incorrect values. In addition to being a powerful tool during model development, tests of anomalous behavior may contribute convincingly to model validity. For example, if a model behaves well except for, say, a limited period of time, and no faults can be found in the model, the error may lie in the data with which the model behavior is being compared. Or, matching the real system's purported behavior may require the inclusion of implausible structure, or parameter values, in the model. In the heroin model, the import data for 1 year were revised downward, because the consumption required to match that import level could be achieved only by an unrealistic increase in heroin user population. Whether due to faults in the model or in the real system, the resolution of the discrepancies found through the anomalous behavior test bolsters confidence and validity.

Behavior Sensitivity Test

Most, but certainly not all, social systems are stable--bureaucracies, in particular, are frequently lampooned for their very, very stable behavior. Small, reasonable changes in a model's parameter values, then, should normally not produce radical behavior changes. If the model's behavior is not seriously affected by plausible parameter variations, confidence in the model is increased. On the other hand, dynamic simulation models are often used to search for parameters that <u>can</u> effect behavior changes. The criterion in the sensitivity test is that any sensitivity exhibited by the model should not only be plausible, but also consistent with observed, or likely, behavior in the real system.

Behavior Prediction Test

Dynamic simulation models are especially useful in predicting how a system would behave if various policies of interest were implemented. Dynamic simulation models offer significant advantages when used in this role; they provide a consistent basis for the predictions. This basis is a consolidation of judgment, experience, and intuition that has been tested against historical evidence, and the predicted effects of implementing alternative policies are promptly available. Confidence in the model is reinforced if the model not only replicates long-term historical behavior, but also responds similarly to existing systems in which various policies have been implemented. For example, over the years many treatment policies have been followed in drug abuse treatment centers. A generic model of such a system, tailored to match any particular center of interest, should replicate the effects produced by the policies implemented in that center.

Family Member Test

Dynamic simulation models acquire added value and confidence when they are generic, i.e., applicable to a family of similar situations, as in the case of treatment centers mentioned above. Drug abuse treatment centers have common basic features, so any one facility may be thought of as a particular case of the basic model embodying these common features. The same is true of payroll, retirement, university, village, city, region, and many other social systems or organizations.

Under these conditions, confidence is enhanced not only because the complementary systems can contribute to the robustness of the model developed for a particular member of the family, but also because the differences among the members can be explicitly identified and defined.

Some family member applications of the heroin model, for example, are readily apparent. The structure is equally applicable to subdivisions of the United States, such as regions or cities. Further, it appears that it is also directly applicable for cocaine, and, possibly, other illegal drug systems.

Behavioral Boundary Test

Exploiting generic models, behavior prediction, and tests of extreme policies may impinge on the model boundary. Is the boundary still adequate for excursions that may extend beyond the region of operation initially envisioned for the model? In prediction, for example, the basic model may have to be revised, so that policy alternatives, or events, such as the impact of discoveries in research programs, can be introduced. In drug abuse models, the inclusion of social trends, or new domestic or foreign policies, may require boundary modifications. The behavioral boundary test is an important step in determining whether the model includes the necessary modifications.

OTHER TESTS

A third class of test--policy implication tests--which includes system improvement, changed behavior prediction, boundary adequacy, and policy sensitivity tests, deals with whether a real system's response to a policy change would replicate the response to the policy change predicted by a model. These tests reflect a different perspective in the application of some of the tests discussed earlier. For example, if real system behavior improves as predicted when tested in a model, was the policy change responsible for the improvement, or were other factors responsible?

This test builds confidence only after numerous real life tests have been completed. The boundary question is inverted: how would boundary changes alter the evaluation of policies and the selection of policies for implementation? These tests tend to be long term and to contribute to confidence and validity most importantly by enlarging the scope of congruence between dynamic simulation models and the systems they represent.

Checking the dimensional consistency of model equations is an additional structural test that may be ignored as trivial, or obvious, but at some peril. For example, if a model contains an equation with heroin expressed in grams on the left side of the equal sign, heroin in grams, and heroin in grams only, must fall out from the right side of the equation. Errors in dimensional consistency can easily creep into model equations during model development and, subsequently, during revisions.

Additional behavioral tests, surprise behavior, and extreme policy behavior can also contribute to confidence and validity. Surprise behavior relates to the recognition of behavior in the real system that was there all along, but not noticed until the system was modeled. Because of its emphasis on identifying the causes underlying observed behavior, System Dynamics readily leads to such discoveries. For example, in the heroin system model, such a surprise was the identification of the relative abundance of heroin as a key parameter influencing the purity, price, and heroin-related deaths that occurred in the real system. In retrospect, like many inventions and discoveries, the relationships may seem very obvious. Such new-found perspectives, of course, contribute significantly to confidence in the model. Extreme policy tests introduce radical policies into the model to see if the behavior of the model is consistent with what would be expected under these conditions. This helps affirm the model's robustness.

COMMON TESTS NOT USED

Paralleling the development of the tests described above has been a growing body of evidence and opinion that many tests commonly associated with model testing are inappropriate, inadequate, or even dysfunctional. In part, these changes derive from the philosophy underlying the System Dynamics method of dynamic simulation modeling, particularly, the notion that all important factors in the real system exerting an influence on the behavior of the system must appear in the model--whether these factors are normally modeled or not. Further, all factors in the model must have a counterpart in the real system. Together with the dynamic, rather than static, nature of the simulation, these characteristics have shifted emphasis from more traditional, statistical tests to the kinds of tests described in this paper--whole model tests that engage all the model variables and their relationships in the testing process.

The t-test, for example, has been shown to be of little use, and possibly misleading, in several studies (Johnson 1980, Mass and Senge 1978).

Briefly, the tests can lead to the exclusion of factors that are important to a model's behavior. Although the tests may be helpful in detecting structural flaws, they are insufficient in the absence of whole model tests. Recently, statistical tests employing Kalman filtering principles have been developed. These tests may be more useful in the development of dynamic simulation models (Peterson 1979). The greater power of these tests stems from their ability to eliminate the effects of measurement error in hypothesis testing.

REFERENCES

Bell, J.A., and Senge, P.M. Enhancing refutability. In: Legasto, A.A. Jr.; Forrester, J.W.; and Lyneis, T.M., eds. System Dynamics. New York: Elsevier North-Holland, 1980. pp. 61-73.

Forrester, J.W. Industrial Dynamics. Cambridge, Mass.: MIT Press, 1961. 464 pp.

Forrester, J.W. Collected Papers of Jay W. Forrester. Cambridge, Mass.: MIT Press, 1975. 284 pp.

Forrester, J.W., and Senge, P.M. Tests for building confidence in System Dynamics models. In: Legasto, A.A. Jr.; Forrester, J.W.; and Lyneis, T.M., eds. System Dynamics. New York: Elsevier North-Holland, 1980. pp. 209-228.

Johnson, C.B. Some effects of data error on econometric models. In: Legasto, A.A. Jr.; Forrester, J.W.; and Lyneis, T.M., eds. System Dynamics. New York: Elsevier North-Holland, 1980. pp. 143-159.

Mass, N.J., and Senge, P.M. Alternative tests for the selection of model variables. IEEE Systems, Man, and Cybernetics, 8(6):450-459, 1978.

Peterson, D.W. Statistical tools for System Dynamics. In: Randers, J., ed. Element of the System Dynamics Method. Cambridge, Mass.: MIT Press, 1979. pp. 224-245.

Roberts, E.B., ed. Managerial Applications of System Dynamics. Cambridge, Mass.: MIT Press, 1978. 699 pp.

AUTHOR

Raymond C. Shreckengost
Central Intelligence Agency
Washington, DC 20505

TELEPHONE SURVEYING FOR DRUG ABUSE: METHODOLOGICAL ISSUES AND AN APPLICATION

Blanche Frank

INTRODUCTION

In the past decade telephone surveying has grown in popularity. Not only have marketing firms used this medium, but so have government agencies, notably the National Center for Health Statistics and the Census Bureau. Although telephone interviewing has limitations, its advantages are making it the dominant method of survey research.

This paper highlights methodological issues in the use of telephone surveys, generally, and for drug abuse, specifically. First, some major issues in telephone surveying are discussed, including sampling, questionnaire design, response rates, data validity, and the management of research using this mode of administration. Then, a New York State telephone survey of drug abuse is described, with emphasis on these methodological issues.

GENERAL METHODOLOGICAL ISSUES

Sampling

A major concern in telephone surveying is the exclusion of non-telephone households. In 1936, when the Literary Digest's telephone survey erroneously predicted that Alf Landon would defeat Franklin D. Roosevelt in the presidential election--at a time when 35 percent of the households had telephones--it was not surprising that much bias was introduced in the sampling (Dillman 1978). By 1981, however, 97 percent of the households in America had telephone service (Census Bureau 1982). Nevertheless, the 3 percent of the households that do not have telephones are surely of interest. They are more likely to be in the South and West, black than white, and in non-Standard Metropolitan Statistical Areas (SMSA) and rural areas (Tyebjee 1979). Some researchers found that households with telephones available were more likely to have "white, male heads of higher average age, income and educational level and to have the spouse present than those households with no telephone available" (Tull and Albaum 1977, p. 394). Thus, given the concentration of these characteristics in the population of interest, a sample can under-represent certain population segments. Some researchers have circumvented this bias by augmenting the telephone sample with personal interviews in census tracts with low telephone penetration.

Sampling for telephone households, however, has essentially relied on two strategies: telephone directories and random digit dialing (RDD). Both have advantages and limitations. Telephone directories are convenient listings of sample units from which a household sample may be drawn, minimizing the nonhousehold units. Sources of bias, however, exist in the use of directories. First, they are out-of-date the moment they are issued because directories cannot include numbers issued in the interim. Given the residential mobility of many Americans, this is surely a consideration. The second and more serious objection is that telephone directories do not include unlisted numbers. About 20 percent of all telephone households are unlisted. The rate of unlisted households varies geographically, with the highest in the Pacific and mid-Atlantic regions and lowest in the South. Urban households with younger heads, fewer children over 12 years old, and more persons between 18 and 34 years old are likely to have unlisted numbers (Tyebjee 1979).

RDD, on the other hand, is a strategy that avoids the sampling biases of telephone directories, and has been shown to produce results akin to areal probability sampling (Klecka and Tuchfarber 1978). Nevertheless, RDD has its own problems. These problems concern the high probability of getting a nonworking or nonhousehold number. Unrestricted random sampling for RDD becomes extremely costly because approximately 80 percent of the numbers in the sampling frame are not assigned to households (Waksberg 1978). Most are either unused, assigned to nonhouseholds, or have some technical difficulty.

Some knowledge of the telephone system is essential. First, depending on the geographic areas of interest, a list of appropriate telephone area codes and existing three-digit working central offices is generally available. Second, it is the practice of telephone companies to assign the four remaining digits in clusters. With this information, a multistage sampling scheme is generally employed in RDD for geographic areas of interest. The scheme admits a cluster of numbers if a first try in a selected series yields a household number (Waksberg 1978, Cummings 1979). Of course, once a cluster of numbers is selected, there are bound to be numbers whose eligibility cannot be determined immediately. Followup calls are necessary. For instance, business telephone numbers can be eliminated quickly by making initial calls during the daytime. Checking with the telephone company about numbers suspected of having technical difficulties will probably clear up those problems.

Another consideration with RDD is the chance of selecting a household with more than one telephone number. About 3 percent or more of telephone households have multiple numbers. These households are more likely to be urban and Eastern and to have more teenagers (Glasser and Metzer 1972). This bias can be corrected by asking respondents the number of telephone numbers reaching the household and weighting the responses by the inverse of this number.

In summary, although nontelephone households obviously are not covered in telephone sampling, RDD can ultimately provide coverage of the sampling frame of households with working telephone numbers. Telephone sampling is frequently compared with sampling for personal interviews. Both types involve a degree of noncoverage. According to one researcher, "Even careful face-to-face surveys probably cannot locate 5 percent or more of all households, although not necessarily the same ones" (Sudman 1981, p. 1).

Questionnaire Design

In general, survey researchers believe that most questions asked in personal interviews may be asked on the telephone. In fact, questions on sensitive topics can move almost freely from one mode to the other (Aneshensel et al. 1982; Freeman et al. 1982). Furthermore, the telephone interview need not be shorter than the face-to-face interview. Depending on the salience of the topic to the respondent, telephone interviews have lasted for an hour or more (Sudman 1981).

There are, however, some caveats in item construction. Because visual aids cannot be used unless special arrangements are made, and respondents can remember only a small number of alternatives, questions should not offer more than three or four alternatives. Questions asking for "yes-no" responses and using branching techniques should be used. Numerical scales are probably preferable to Likert-type scales.

Some items of interest have shown more respondent resistance than others on the telephone. Income is one such item. Many respondents are particularly suspicious when income questions are asked. These are often placed at the end of the interview after there has been an opportunity to build rapport. Race is another item that sometimes meets resistance. This question should be asked directly on the phone or omitted. In any case, a simple but meaningful introduction is important to "grab" the respondent in a telephone interview.

Response Rates

In discussing response rates, it is important to distinguish between refusals and "not-at-homes." Both are extremely important in probability surveys.

In general, telephone surveying has recorded a higher refusal rate than face-to-face interviewing. The refusal rates in national telephone surveys, for instance, are at least 5 percentage points higher than those recorded for personal interviews (Groves and Kahn 1979). A variety of factors may contribute to refusals, such as characteristics of the respondents, characteristics of the interviewer that are detectable in speech, and the conditions of the interview, such as the time of the day.

Little investigation has been conducted into these factors, although some findings indicate that refusals are more likely to occur among the poorest households, underprivileged minorities, households with an older head, and households with a head having little education (O'Neil 1979). A recent study has shown some interesting interviewer effects on response rates. When interviewing experience is controlled, older interviewers and interviewers with optimistic expectations achieved higher response rates (Singer et al. 1983). In addition, a study found that the timing of the call affected response rates. Refusals were higher on weekends than weekdays, highest in the evenings, and lowest in the mornings (Falthzik 1972). These factors in telephone surveying should be explored further.

Several techniques, however, have been used to minimize refusals. One technique referred to as "foot-in-the-door" relies on a mini-interview of five questions to get initial compliance and to schedule a longer interview later. This technique used in an experimental group received a higher response rate than the control group, where the technique was not used (Reingen and Kernan 1977).

Another technique used mixed-mode surveying. The telephone respondents who refused cooperation were followed up with a self-administered mail questionnaire, a personal interview, or both to obtain cooperation. Response rates ultimately increased to about 90 percent (Siemiatycki 1979).

An interesting observation about refusals is that telephone screening methods can reach respondents who are more difficult to locate in face-to-face interviewing, but obtain a lower response rate from respondents who are located, i.e., those who reside in security-conscious, high-rise apartments (Sudman 1981).

The second source of nonresponse--the not-at-homes--can create a systematic bias in the sample. Empirical evidence indicates that those at home are overrepresented by respondents in the over-64-years age group, those with low education and low income, those with home-related occupations, and those who reside in rural places (Dunkelberg and Day 1973).

There are essentially two ways of dealing with the not-at-home nonresponse problem. One is to use a correction factor that weights each completed interview by the inverse of the probability of finding the respondent at home, such as the Politz-Simmons correction. The factor is derived from a question that asks, for instance, on how many nights in the past week the respondent was at home (Politz and Simmons 1949).

Another approach to minimizing this problem is to employ a rigorous policy of call-backs at various times during the day and week. Although a policy of three or four call-backs is often used, residents of larger SMSAs--especially in the Northeast--require more than four call-backs, compared with two in smaller areas (Tyebjee 1979).

Validity of Data Collected

The literature is replete with studies that find results for telephone surveying not significantly different than data collected from personal interviewing (Coombs and Freedman 1964; Hochstim 1967; Klecka and Tuchfarber 1978). Many of these studies have randomly assigned respondents to one mode of interviewing or the other, and then compared the findings. Nevertheless, as already indicated, different interview methods tend to cover somewhat different segments of the population, data on income are not easily shared over the telephone, and telephone interviewers may affect outcomes in ways that are not entirely understood.

Some evidence is offered that respondents do not put as much effort into a telephone interview as they do into a face-to-face interview (Groves and Kahn 1979). Phone interviews are generally shorter, and studies have found a lack of the richness of data collected by other modes of administration (Siemiatycki 1979).

Management of Telephone Surveying

Some of the major advantages of telephone surveying come from management considerations such as cost, quality control, and time. As far as cost is concerned, it is generally held that telephone interviewing costs about 50 percent or less of the cost of personal interviewing (Siemiatycki 1979; Sudman 1981). Given the centralization of telephone facilities and the computerization of many procedures in current telephone surveying compared to travel costs and data processing costs in personal surveying, it is not difficult to understand the economies. According to one researcher, more than half of the costs of face-to-face interviewing involve interviewer travel to locate respondents (Sudman 1981).

A second consideration is the quality control that can be exercised in the several steps of the surveying process. With centralized facilities, telephone interviewing can be monitored easily and problems can be detected very early in the data collection. Furthermore, the use of computer-assisted telephone interviewing, known as CATI, minimizes error on the part of the interviewer and inconsistent responses on the part of the respondent. It allows the use of a complicated questionnaire design with elaborate skip instructions and forward feedback instructions.

Finally, analyses may be performed even before all the interviews are completed. For instance, analyses of respondents' profiles can indicate whether subsequent sampling may have to concentrate on underrepresented segments of the population (Tyebjee 1979).

Time is also an important consideration. The interval between data collection and the reporting of findings is shortened since steps in the process such as coding, keypunching, and verifying are eliminated in CATI surveys. Soon after the last interview is completed, a computer tape of responses checked for reliability is ready for analysis.

AN APPLICATION OF TELEPHONE SURVEYING IN A STUDY OF DRUG ABUSE

Given these general methodological considerations in telephone surveying, how do they come into play in a household survey of medical, non-medical, and illicit substance use? The remaining discussion deals with New York State's experience with telephone surveying.

The New York State Division of Substance Abuse Services periodically conducts a household survey of drug use across the State. The purpose of the survey is to monitor the drug abuse problem among household residents and to estimate the number of drug abusers in this population. With the rising cost of face-to-face interviewing and the reluctance of respondents to open their doors--especially in New York City--a telephone survey seemed a probable alternative to a face-to-face survey. In order to determine the feasibility of this mode of administration in a survey of sensitive, stigmatizing, and illicit behavior such as drug abuse, it was decided to conduct a pilot study first. In 1980, the pilot study was conducted. The results did show that such a survey was feasible. In 1981, a full-scale computer-assisted telephone survey was conducted in the State. The discussion that follows highlights the methodological issues and the ways in which these were handled in the pilot telephone survey and the full-scale survey.

Sampling

New York State's telephone survey tried sampling nontelephone residents as well as telephone households. The nontelephone segment included the residents of Single Room Occupancy (SRO) hotels in New York City. This subgroup had never been included in a probability sample studying drug abuse, and it is a population that has been of particular interest in the city. The approximately 200 SRO hotels with their 23,000 occupants provided a fairly well-defined frame from which residents could be sampled.

Based on experience gathered from the pilot study, a sampling scheme was worked out. A sample of 47 SRO hotels was selected from the universe, which was stratified by borough and price range. Each building manager of a selected hotel received a letter explaining the study and requesting cooperation. Each manager was paid $25 for assistance in sharing information about the number of rooms, the layout of the hotel, and the availability of a pay telephone. State field workers were able to sample the rooms systematically with a different random start for each hotel. The field workers would knock at the door of a selected room and try to elicit cooperation from the resident. If and when the resident agreed, he/she was walked to the phone, a phone call was made to the central telephone facility, and the respondent was interviewed out of earshot of the fieldworker. When the respondent said the interview was completed, the fieldworker would speak to the interviewer to determine whether the interview was in fact completed. At that time the respondent received $10 for his/her time and effort. This combination of field worker and telephone interviewing yielded a sample of 236 respondents from 43 SRO hotels.

The sampling of the telephone households eventually used RDD. During the pilot study, however, the use of the telephone directory was also attempted with an advance letter sent to the selected households explaining the study, stating that a call would be made to them at a future date, and offering $10 for completing the interview. The check could be donated to a charity or mailed directly to the respondent. An attempt was also made with a similar advance letter that asked the respondent to call the contractor using an 800 toll-free telephone number. Response rates based on completed interviews as a proportion of eligible households contacted were calculated for each strategy. As the pilot study turned out, the RDD sample yielded a response rate of 69 percent, the advance letter indicating that the contractor would call yielded 72 percent, and the advance letter asking the respondent to call yielded a 9 percent response rate. Since the response rates for the RDD sample and the telephone directory sample with the first advance letter were not significantly different, and since the RDD sample assures inclusion of nonlisted households and the ability to maintain the anonymity of the respondent, the RDD strategy for sample generation was used for the full-scale study.

Although RDD can be a costly effort, the accumulated knowledge of the telephone research organization allowed efficient sampling in the multistage design. First, the research organization had available a listing for each of the approximately 2,000 telephone exchanges or central offices in use in New York State. Each exchange with its six-digit identifier had associated geographic information, including county, and city within county, which was important to the stratification design used in the study. A strictly proportional sample of exchanges was selected to represent the metropolitan and nonmetropolitan counties in each region of the State. Then, using a random digit generating formula calculated by the computer, an equal or known number of four-digit combinations was selected within each central office. Using the cluster sampling described earlier, a sample was generated. Since a larger proportion of the nonworking and nonhousehold numbers in each telephone exchange was identified using prior knowledge, approximately 70 percent of these numbers were identified and eliminated before the interviewing got started. This hurried the sampling process along and cut costs. An RDD-generated sample of 3,251 householders participated in the survey, representing the demographic and educational characteristics of the total household population as well as their metropolitan and nonmetropolitan location in the State.

An effort was also made in the pilot study to sample two "rare trait" subgroups: American Indians and high school dropouts. The literature does speak of RDD as useful in locating rare populations, especially if a large sample is generated (Waksberg 1978). Using network sampling--where respondents were asked to provide names of such individuals--for both subgroups, and the telephone directory for surnames that might be native American, some interviews were conducted. Nevertheless, this type of network and telephone directory sampling for subgroups was not pursued in the full-scale study.

Questionnaire

The questionnaire included numerous items on the use of prescription drugs and illicit substances, the possible consequences of that drug use in terms of problems encountered in everyday life, self-perception of drug dependence, and the need for treatment. Despite the sensitive nature of many questions, the number of respondents having any knowledge of the subject matter who refused to complete the interview was minimal.

Some of the drawbacks considered earlier in questionnaire design did not present a problem in the survey. For instance, in face-to-face drug use surveys, visual aids are often used to help respondents identify prescription drugs that they have used. In a telephone survey, however, having such stimuli available would require some prearrangements. The contractor did suggest that if telephone directory sampling was used, a lockbox might be mailed to each household prior to the telephone interview. When the respondent is reached on the phone, the combination to open the lockbox is given, and within the box would be the visual aids required. The curiosity aroused not only increases the interest in the interview but provides the necessary materials. Since the use of directories was not planned and the respondents' anonymity was important, this device was not used. Instead, to help the respondents remember the names of the drugs they used, the interviewer had a comprehensive list of drug names by category of drug programmed into the computer. This list was referred to when questions of drug identification arose.

The household income question yielded better results than expected. The question was among the last in the interview and the missing data were minimal. Recently, the findings for drug use by income were analyzed. The findings showed that the distribution of the weighted sample by household income was almost identical to the 1980 Census distribution of New York State households by family income.

Response Rates

To encourage participation, respondents were offered compensation for their time. Respondents in New York City were offered $10 for completing the interview, and $5 was offered in the rest of the State. This money could be sent to the respondent, another individual selected by the respondent, or to a charity of the respondent's choice. Most respondents opted to receive the money.

The response rate for the RDD portion of the survey was 66 percent--a rate similar to the face-to-face administered survey the State conducted in 1976. In the telephone survey, there were 3,251 completed interviews of 4,956 eligible households. These eligible households include refusals where eligibility was established (429) and refusals where eligibility was not established (estimated 1,276). It was assumed from prior experience that 74.3 percent of the refusals with undetermined eligibility were in fact eligible. The calculation excludes troublesome categories that could not be resolved despite five callbacks and some nonresponse followup where refusals were not received. These categories include: no answer/busy (829), callbacks (263), and "language barriers"(454).

What contributed to the refusal rate was the need to oversample in five central cities--where refusal rates were the highest--and to oversample youth between the ages of 12 and 17 years, who required parental permission for participation.

Validity of the Data

One of the purposes of the pilot study was to determine whether the telephone as a mode of data collection could yield valid rates of drug use. The findings from the 1980 pilot study were compared to rates obtained in the 1979 National Household Survey conducted by the National Institute on Drug Abuse (Fishburne et al. 1979). The rates proved very similar. Compared to New York State's 1976 household survey, the rates for the 1980 pilot study and the 1981 full-scale study were higher for several drugs, which reflected trend data from a variety of indirect indicators.

A question about validity, however, may pertain to the findings for respondents 12 to 17 years of age. Since these respondents needed parental permission for participation, it was entirely possible that someone else was present during the interview, interfering with the candor of responses. In general, this obstacle to participation may have acted as a biasing factor. Compared to self-administered school survey findings for this age group, the telephone survey found very low rates of drug use.

Perhaps the most interesting experience in this project was the attempt to improve the validity of the findings by use of the randomized response technique on the telephone. This technique was first developed in the mid-60s to enhance the privacy of the respondent's answers to sensitive items (Warner 1965). The technique usually involved a randomizing device, such as a pair of dice, and two statements--one about a sensitive attribute, such as, "I have used heroin," and one about an innocuous attribute, such as, "I have watched television." Based on some rules, the respondent uses the randomizing device to determine the statement to which the "Yes" or "No" response is given. The interviewer simply records the response without knowing the statement to which it applies. Based on the probabilities associated with the results of the game, and the assumption that the respondent is playing according to the rules, estimators of the proportion of the population having the sensitive characteristics may be determined. This technique, however, was designed for use in personal interview surveys. Thus, its use on the telephone in the pilot study was a new undertaking for the contractor.

After much discussion, two variations of the randomized response technique were incorporated into the interview. First, rather than two statements, one question was employed so that the respondent would not have to remember too much. Second, the randomizing device was something that was readily available in the home with known probability that would optimize the percentage of response based on the truth. The randomizing device was three coins. Each of the selected respondents was asked to toss the coins before each question and to answer based on certain rules. Several questions using this randomized response technique were asked about each of four drugs: cocaine, LSD, Angel Dust, and heroin.

Of the sample of 203 household respondents in the pilot study, 115 were randomly selected for randomized response and 88 were asked the questions directly. Of the 115, only 60 agreed to use the technique. Of the 55 who rejected its use, 33 claimed they never used the drugs and it would be a waste of time, and 22 wanted to tell the truth without playing games. In general, there were no significant differences in findings between the 88 respondents originally selected to answer by direct questioning and the 66 who participated in the randomized response technique.

This technique was not used in the full-scale study for several reasons. The technique, by definition, prevents knowledge of individual behavior and only allows findings for the group. Thus, this drawback inhibits the data analyses that may be performed. Second, the technique reduces the effective size of the sample and thus the power of the statistical techniques. Third, the additional complexity increases the length of the interview, which has implications for cost. Finally, it became clear from the pilot study that many respondents did not want to be bothered with the technique. On the other hand, certain people seemed to be at ease with it; these were mainly the younger adults. Although our experience indicates that the technique may not be suitable for a total sample, it might be offered to those who feel uneasy about answering certain questions. Nevertheless, the telephone turned out to be a more flexible medium than had been expected.

Management of the Survey

Some of the obvious advantages of the telephone survey were the computer-assisted capabilities, the ease in monitoring the interview, and the timeliness of the process. The use of the computer in programming the questionnaire and its complex skip patterns, in customizing the questionnaire to reflect answers already given, in recording the responses, and in flagging inconsistencies minimized interviewer error and editing problems. A clean data tape was delivered soon after the last interview was completed.

Being able to monitor the interviews was a reassuring experience. It became clear that sensitive drug use questions could be asked on the telephone, and that interviewers were very adept in asking the questions, including the use of the randomized response procedures in the pilot study. The ease in supervision surely enhanced the quality of interviewing.

Finally, the timeliness was most remarkable. Data collection started in the beginning of March 1981; the first report of prevalence and incidence was issued in September 1981. Six months was the fastest turnaround for any of the population surveys this agency had conducted.

SUMMARY AND CONCLUSIONS

In light of New York State's experience, it is probable that future household drug use surveys will use telephone administration. Drug use questions are not as sensitive as had been thought, and are easily administered by telephone. In addition, the lower costs, the computer-assisted capabilities, and the saving in time are some of the advantages in comparison to face-to-face surveying. In order to address the nontelephone segments of the household population--despite their declining proportion--and to improve response rates, mixed-mode interviewing may have to be considered. Given a better understanding of telephone-associated behavior and the increasing popularity of technological advances, such as the portability and mobility of phones, telephone surveying may become even more attractive in the future.

REFERENCES

Aneshensel, C.S.; Frerichs, R.R.; Clarke, V.A.; and Yokopenic, P. Measuring depression in the community: A comparison of telephone and personal interviews. Pub Opin Q, 46:110-121, 1982.

Coombs, L., and Freedman, R. Use of telephone interview in a longitudinal fertility study. Public Opin Q, 28:112-117, 1964.

Cummings, K.M. Random digit dialing: A sampling technique for telephone surveys. Public Opin Q, 43:233-244, 1979.

Dillman, D.A. Mail and Telephone Surveys: The Total Design Method. New York: John Wiley & Sons, 1978. 325 pp.

Dunkelberg, W.C., and Day, G.S. Non-response bias and call-backs in sample surveys. J Market Res, 10:160-168, 1973.

Falthzik, A. When to make telephone interviews. J Market Res, 9:451-452, 1972.

Fishburne, P.; Abelson, H.; and Cisin, I. National Survey on Drug Abuse: Main Findings, 1979. National Institute on Drug Abuse. Washington, D.C.: U.S. Govt. Print. Off., 1980. 169 pp.

Freeman, H.E.; Kiecolt, K.J.; Nichols, W.L. II; and Shanks, J.M. Telephone sampling bias in surveying diability. Public Opin Q, 46:392-407, 1982.

Glassner, G., and Metzger, G.D. Random digit dialing as a method of telephone sampling. J Market Res, 9:59-64, 1972.

Groves, R.M., and Kahn, R.L. Surveys by Telephone: A National Comparison With Personal Interviews. New York: Academic Press, 1979. 358 pp.

Hochstim, J.R. A critical comparison of three strategies of collecting data from households. J Am Stat Assoc, 62:976-989, 1967.

Klecka, W.R., and Tuchfarber, A.J. Random digit dialing: A comparison to personal surveys. Public Opin Q, 42:105-114, 1978.

O'Neil, M.J. Estimating the nonresponse bias due to refusals in telephone surveys. Public Opin Q, 43:218-232, 1979.

Politz, A., and Simmons, W. An attempt to get the not-at-homes into the sample without call-backs. J Am Stat Assoc, 44:9-31, 1949.

Reingen, P.H., and Kernan, J.B. Compliance with an interview request: A foot-in-the-door, self-perception interpretation. J Market Res, 14:365-369, 1977.

Siemiatycki, J. A comparison of mail, telephone, and home interview strategies for household health surveys. Am J Public Health, 69(3):238-245, 1979.

Singer, E.; Frankel, M.R.; and Glassman, M. The effect of interviewer characteristics and expectations on response. Public Opin Q, 47:68-83, 1983.

Sudman, S. Telephone methods in survey research: The state of the art. Paper read at American Educational Research Association Meeting, 1981, Los Angeles.

Tull, D.S., and Albaum, G.S. Bias in random digit dialed surveys. Public Opin Q, 41:389-395, 1977.

Tyebjee, T.T. Telephone survey methods: The state of the art. J Market Res, 43:68-78, 1979.

U.S. Bureau of the Census. Statistical Abstract of the United States: 1982-83. Washington, D.C.: U.S. Govt. Print. Off., 1982. 1008 pp.
Waksberg, J. Sampling methods for random digit dialing. J Am Stat Assoc, 73:40-46, 1978.
Warner, S.L. Randomized response: A survey technique for eliminating evasive answer bias. J Am Stat Assoc, 60:63-69, 1965.

AUTHOR

Blanche Frank, Ph.D.
New York State Division of Substance Abuse Services
2 World Trade Center
New York, NY 10047

A PILOT STUDY ASSESSING MATERNAL MARIJUANA USE BY URINE ASSAY DURING PREGNANCY

Barry S. Zuckerman, Ralph W. Hingson, Suzette Morelock, Hortensia Amaro, Deborah Frank, James R. Sorenson, Herbert L. Kayne, Ralph Timperi

INTRODUCTION

Approximately 10 percent of childbearing-age women in the United States smoke marijuana (Fried 1980). Studies of both middle- and low-income women in the Boston area indicate that 10 to 15 percent of pregnant women acknowledge smoking marijuana during pregnancy (Hingson et al. 1982; Wilner 1981). Because many women do not realize they are pregnant until at least 1 or 2 months into their pregnancy, it is possible that the proportion of women who consume marijuana during the first trimester of pregnancy may be even higher.

The main ingredient of marijuana, delta-9-THC, is known to cross the placental barrier; in early pregnancy the transfer is higher than in late pregnancy (Harbison and Mantillaplata 1972). Smoking marijuana during pregnancy raises the possibility of fetal toxicity through placental transfer.

Some studies with animals demonstrate an association between marijuana exposure and intrauterine growth retardation (Abel 1980). Conclusions from these studies are limited because of lack of an inhalation model for animals and lack of pair-fed controls.

Studies on humans also demonstrate conflicting findings about the relationship between marijuana use during pregnancy and adverse perinatal outcome. Associations have been observed between marijuana use and a greater likelihood of meconium staining and precipitative labor (Greenland et al. 1982). Another study (Hingson et al. 1982) demonstrated that marijuana use was independently associated with lower birthweight when other maternal characteristics that might influence fetal growth were controlled. Results of that study indicated that women who used marijuana fewer than three times per week during pregnancy delivered infants who were 95 grams smaller than infants of nonusers. Women who used marijuana three or more times per week delivered babies 139 grams smaller than infants of nonusers. In comparison, women who smoked one pack or more of cigarettes daily delivered infants who were 83 grams smaller than infants of nonusers. Another study (Linn et al. 1983) did not demonstrate an association between marijuana use during pregnancy and low birthweight when possible confounding variables were controlled.

Possible explanations for these conflicting findings include the following: 1) the different studies represent different populations of pregnant women; 2) the dependent variables of growth were assessed differently (record review vs. examination); and 3) the independent variable of drug use was assessed differently (prospective vs. retrospective self-report).

A prospective study using urine assays to detect the presence of marijuana has been designed to overcome some of the limitations in the previous studies, including our own. A prospective design should diminish recall bias of self-reported drug use. The urine assay will be used to confirm the validity of reported marijuana use. The main hypothesis of the prospective study, to be conducted by the authors, is: When potentially confounding factors are analytically controlled, mothers who smoke marijuana during pregnancy will deliver infants who are significantly smaller, exhibit more congenital anomalies, and demonstrate greater neurobehavioral dysfunction than infants of mothers who do not smoke marijuana. An important component of the study will be the measurement of marijuana in urine during pregnancy. This information will supplement interview results and will address the inherent difficulty in obtaining valid and reliable interview data.

Because the proposed study includes the novel feature of requesting a prenatal urine sample to assess drug use, a pilot study was needed to determine whether these women would be willing to participate.

This report presents the results from the pilot study. The aims of this pilot study were:

1) to assess study participation rates among pregnant women informed that their drug consumption would be assessed by urine assay;

2) to compare rates of marijuana consumption as determined by self-report versus urine assay among participating women;

3) to obtain more contemporary data than in our 1977-79 sample (Hingson et al. 1982) on the levels of marijuana, alcohol, psychoactive drug, and cigarette use during pregnancy among women who receive prenatal care at Boston City Hospital (BCH).

PROCEDURE

Women registering for prenatal care at BCH Prenatal Clinic from June 13, 1983 through July 14, 1983 were requested to participate in the study. The study was conducted on 16 of the 21 clinic days during the study period. All pregnant women who spoke English or Spanish were eligible.

Interviews were conducted by a bilingual female psychologist. Data on the frequency and quantity of marijuana use as well as a variety of other health habits and characteristics thought to influence fetal development were gathered through a close-ended questionnaire. The interview time ranged from 20 to 60 minutes. Women were asked whether they had ever used marijuana, and, if so, whether they had used it during pregnancy. Women who reported using marijuana during pregnancy were asked about the frequency of use during each trimester and during the past week. This last time period was requested in order to compare self-reported findings to the results of the urine assay. The 1-week period represents a conservative estimate of the assay sensitivity.

During the pilot month, 269 women visited the prenatal clinic. Our interviewer asked 81 of those women to participate. Selection of subjects was based on interviewer availability. Those asked to participate did not differ in age, race, marijuana use, or other drug use, as reported to the clinic staff, from those not asked to participate.

Upon arrival at the interviewer's office, women were informed of the nature of the study and provided with both a verbal description and written informed consent form that described the study in detail and the use of both interview and urine samples in detail. They were also informed that they would be paid $5.00 for their time. The interviews were conducted immediately after consent was obtained.

After the interview, participants provided a urine specimen, which was immediately labeled with a subject identification number and refrigerated in the clinic. At the end of each day, urine specimens were sealed inside foam coolers filled with ice packets. Urine samples remained refrigerated inside these coolers for 12-24 hours (depending on the time of collection) before being collected by the clinic courier service for transportation to the Massachusetts State Laboratory.

Seventy-five urine samples were received in the laboratory and frozen at -20 degrees C in plastic containers until analyses were performed. Detection of cannabinoid metabolites in urine was done using the enzyme-mediated immunoassay technique (EMIT) (Rowley et al. 1976). Three known calibrator samples were used in the assays: a negative calibrator containing no metabolite (0.0 mg/ml), a low calibrator containing 20 mg/ml of the 11-nor-delta^{8}-THC-9 carboxylic acid derivative of delta^{9}-THC, and a medium calibrator containing 75 mg/ml of the derivative. The low calibrator (20 mg/ml) is used as the cutoff value for positive/negative interpretation and ensures at least 95 percent confidence in the positive/negative classifications (DeLaurentis et al. 1982). Levels of urinary metabolites are detectable within a few hours after exposure to marijuana (Rodgers et al. 1978) and remain detectable 7 to 10 days after smoking (Clark et al. 1980). Frequent users often have continually detectable baseline levels. Usually as much as 50 percent of an initial dose is excreted within 72 hours.

Samples were run in duplicate. Positive samples were tested again in a separate assay, and distilled water blanks were used as spacers in this repeat assay to ensure against the possibility of carryover. Confirmation of 13/18 EMIT THC-positive urine samples was done by gas chromatography (GC) (Whiting et al. 1982). There was an insufficient volume of urine to do a confirmatory analysis by GC for five samples. The confirmatory procedure only measures the primary urinary metabolite of THC and would be expected to agree with EMIT THC-positive results in 90 percent of the samples.

RESULTS

The results will be presented in the order of the goals of the pilot study.

Of the 81 pregnant women asked to participate, 6 (7 percent) refused. Three of them refused because they did not have adequate uninterrupted time to stay for the interview. Following establishment of a uniform procedure for interviewing women during the second week of the study, there were no refusals for this particular reason. To our knowledge, no women cited the urine assay explicitly as a reason for nonparticipation. The refusal rate of 7 percent compares favorably with that in our previous study (14 percent) (Hingson et al. 1982) and suggests that women are willing to participate in a study that tests their urine for marijuana and psychoactive drugs.

Of the study sample of 75 women, 38 (51 percent) reported smoking marijuana at some time in their lives. Of these, 23 women (31 percent) reported smoking marijuana during their pregnancy, and 11 women (15 percent) reported smoking in the week prior to the interview (table 1).

TABLE 1. *Self-reported marijuana use*

Marijuana Use	Number of Women (%) (N=75)
Never smoked marijuana	37 (49%)
Ever smoked marijuana:	38 (51%)
Prior to but not during pregnancy	15 (20%)
During pregnancy	23 (31%)
During week prior to interview	11 (15%)

Of the 75 women in the study, 18 (24 percent) had urine samples that were positive for the presence of marijuana metabolites (table 2). Sixteen of the positive samples had greater than 75 mg/ml of delta^9-THC metabolite when assayed by the EMIT method. The other two EMIT-positive samples had 45 and 60 mg/ml of the metabolite. Ninety-two percent of the EMIT-positive urine tests were confirmed by a blindly conducted gas chromatography test (table 3). This is consistent with rates reported in the literature (CDC 1983). Calculations from the data in table 2 indicate the sensitivity of self-report is 56 percent with a specificity of 98 percent. The low sensitivity focuses on the large numbers of false negatives while the high specificity demonstrates the small number of false positives. Eight women (12.5 percent) reported not having smoked marijuana in the previous week but had positive urine findings. While all eight women reported having used marijuana at some point, four of them reported using marijuana prior to but not during their pregnancy. The other four acknowledged using marijuana during pregnancy, but not during the week before the interview. Had we relied on self-report alone, we would have missed 15 percent (4/27) of the women who used marijuana during pregnancy and 44 percent (8/18) who used it in the previous week. One woman who reported smoking marijuana in the past week had a negative urine test.

TABLE 2. *Marijuana use in the previous week assessed by self-report and cannabinoid assay*

		Cannabinoid Assay		Total
		(+)	(-)	
Self-report	(+)	10	1	11
	(-)	8	56	64
Total		18	57	75

TABLE 3. *Confirmatory tests of urine samples found to contain cannabinoid metabolites by EMIT assay*

Sample Number	EMIT THC	GC RRT*
1	positive	1.29
2	positive	1.29
3	positive	QNS**
4	positive	1.37
5	positive	1.29
6	positive	1.36
7	positive	QNS
8	positive	1.29
9	positive	1.31
10	positive	1.35
11	positive	1.37
12	positive	QNS
13	positive	1.37
14	positive	negative
15	positive	1.29
16	positive	QNS
17	positive	1.31
18	positive	QNS

*Relative retention time (RRT) of THC-COOH compared to oxyphenbutazone by gas chromatography (GC).

**Quantity not sufficient for analysis.

In the pilot sample a greater percentage of pregnant women reported smoking marijuana, smoking cigarettes, and drinking during pregnancy than had reported such behavior during our 1977-79 study at Boston City Hospital (table 4). The knowledge that urine tests would assess the use of these substances in our pilot study may have prompted a greater proportion of women to report such use in our pilot study than in our earlier study, in which urine samples were not collected. However, other possible reasons for the differences can also be hypothesized, e.g., differences in the characteristics of women using the hospital prenatal clinic now compared to 1977-79, changes in maternal habits now compared to the late 1970s, or greater accuracy of data on such habits collected during pregnancy as opposed to after delivery, as in our earlier study.

TABLE 4. *Self-reported substance use any time during pregnancy*

Substance Use During Pregnancy	Boston City Hospital Study Feb. 1977 - Oct. 1979* Percent (N=1,690)	Pilot Project June 1983 Percent (N=75)
Marijuana	14	31
Cigarettes	48	60
Psychoactive Drugs	1	11
Alcohol	38	64

*Hingson et al. 1982.

Whatever the reasons for the higher levels of marijuana use, psychoactive drug use, smoking, and drinking during pregnancy reported in our pilot sample, they suggest an increased ability for the proposed study to detect whether an association exists between maternal marijuana use and adverse fetal development. On the basis of cross-tabulations between self-reported marijuana use and other substance use in our pilot data, we project that in the proposed sample of 1,500 women, 240 will smoke marijuana but not take other psychoactive drugs during pregnancy, 60 will smoke marijuana but not cigarettes, and 40 will smoke marijuana but not drink at all during pregnancy. If marijuana use during pregnancy is identified by either self-report or urine assay, 540 women will be identified as using marijuana. Of these, 320 will not use other psychoactive drugs, 60 will not smoke cigarettes, and 60 will not drink during pregnancy. These numbers are considerably larger than those in our earlier study, in which we were still able to identify an independent association between maternal marijuana use and reduced fetal growth using multiple regression techniques.

The markedly higher proportion of women who reported other psychoactive drug use during pregnancy than in our earlier study (11 percent vs. 1 percent) may, if it persists, also afford us a greater opportunity to assess whether use of other psychoactive substances may relate to adverse fetal development.

DISCUSSION

The testing of urine for marijuana and other drug use resulted in the identification of 15 percent more women who used marijuana during pregnancy than by use of self-report alone. This pilot study reassures us that our proposed prospective study utilizing a urine assay for marijuana use will provide more reliable answers to the questions about the possible effects of maternal marijuana use on human fetal growth than any research conducted to date.

The data from this pilot study further suggest that we will be able to obtain a sufficiently high participation rate to ensure recruitment of a sample of at least 1,500 mothers. The logistical and practical problems of conducting the proposed project are manageable. Finally, we will have an adequate number of cases of individuals with and without marijuana and other drug exposure during pregnancy to assess whether or not marijuana has an independent effect on fetal outcome.

The differences observed in our pilot project in self-reported marijuana use in the week prior to interview and urine assays raise important questions about the validity of using only self-reports of marijuana use to assess the possible effects of maternal substance use during pregnancy on fetal development. One might speculate that in the absence of urine samples, people might systematically underreport drug use and other habits. If the underreported habits are also associated with adequately reported other habits (e.g., nutrition), then the effects of the underreported factors on infant outcome may be underestimated. Whether self-reports consistently produce underreporting of marijuana (and use of other substances as well) should be a fundamental concern for researchers in this area.

The effect of testing urine for drugs on women's self-report of drug use is also an important methodologic question. As we have demonstrated, reported use of all drugs is much higher in the present pilot study than in our previous study, in which there was no urine testing for drug use. We are undertaking a randomized control study evaluating the effect of urine assays on self-report of drug use. Our hypothesis is that women will more often report drug use when they know their urine will be tested. The results of this study will be important to future studies assessing drug use.

REFERENCES

Abel, E.L. Prenatal exposure to cannabis: A critical review of effects on growth development and behavior. Behav and Neural Biol, 29:137, 1980.

Center for Disease Control. Morbidity and Mortality Weekly Report, 32(36):469-471, 1983.

Clark, S.; Turner, J.; and Bastiani, R. EMIT(R) Cannabinoid Assay: Clinical Study No. 74 Summary Report. Palo Alto, CA: Syva, 1980. 28 pp.

DeLaurentis, M.J.; McNeil, K.; Mann, A.J.; Clark, S.; and Greenwood, H.M. An EMIT assay for cannabinoid metabolites in urine. In: Hawks, R.L., ed. The Analysis of Cannabinoids in Biological Fluids. National Institute on Drug Abuse Research Monograph 42. DHHS Pub. No. (ADM) 82-1212. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1982. pp. 69-84.

Fried, P.A. Marijuana use by pregnant women: Neurobehavioral effects in neonates. Drug Alcohol Depend, 6:415-424, 1980.

Greenland, S.; Staish, D.; Brown, N.; and Gross, S. The effects of marijuana use during pregnancy. Am J Obstet Gynecol, 143:408, 1982.

Harbison, R.D., and Mantillaplata, B. Prenatal toxicity, material distribution and placental transfer of tetrahydrocannabinol. J Pharmacol Exp Ther, 180:446, 1972.

Hingson, R.; Alpert, J.; Day, N.; Dooling, E.; Kayne, H.; Morelock, S.; Oppenheimer, E.; and Zuckerman, B. Effects of maternal drinking and marijuana use on fetal growth and development. Pediatrics, 70:539, 1982.

Idanpaan-Haikkila, J. Placental transfer of tritiated I-delta 9 tetrahydrocannabinol. N Eng J Med, 281:330, 1969.

Linn, S.; Schoenbaum, S.C.; Monson, R.R.; Rosner, R.; Stubblefield, P.C.; and Ryan, K.J. The association of marijuana use with outcome at pregnancy. Am J Public Health, 70:1161, 1983.

Rodgers, R.; Crowl, C.P.; Eimstad, W.M.; Hu, M.W.; Kam, J.K.; Ronald, R.C.; Rowley, G.L.; and Ullman, E.F. Homogeneous enzyme immunoassay for cannabinoids in urine. Clin Chem, 24:95-100, 1978.

Rowley, G.L.; Armstrong, T.A.; and Crowl, C.P. Determination of THC and its metabolites by EMIT homogeneous enzyme immunoassay: A summary report. In: Willette, R.E., ed. Cannabinoid Assays in Humans. National Institute on Drug Abuse Research Monograph 7. DHEW Pub. No. (ADM) 78-339. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1976. pp. 28-32.

Whiting, J.D., and Manders, W.W. Confirmation of tetrahydrocannabinol metabolite in urine by gas chromatography. J Anal Toxicol, 6:49-52, 1982.

Wilner, S. A comparison of smoking and other maternal behaviors during pregnancy in a health maintenance organization and fee for services practices. In: Wilner, S. Maternity Care in Two Health Care Systems. Doctoral Dissertation, Harvard University, June, 1981.

AUTHORS

Barry S. Zuckerman, M.D.
Department of Pediatrics
Boston City Hospital
818 Harrison Avenue
Boston, Massachusetts 02118

Ralph W. Hingson, Sc.D.
School of Public Health
Boston University School of Medicine
80 East Concord Street
Boston, Massachusetts 02118

Suzette Morelock, Ed.M.
School of Public Health
Boston University School of Medicine
80 East Concord Street
Boston, Massachusetts 02118

Hortensia Amaro, Ph.D.
School of Public Health and Department of Pediatrics
Boston University School of Medicine
80 East Concord Street
Boston, Massachusetts 02118

Deborah Frank, M.D.
Department of Pediatrics
Boston City Hospital
818 Harrison Avenue
Boston, Massachusetts 02118

James R. Sorenson, Ph.D.
School of Public Health
Boston University School of Medicine
80 East Concord Street
Boston, Massachusetts 02118

Herbert L. Kayne, Ph.D.
School of Public Health
Boston University School of Medicine
80 East Concord Street
Boston, Massachusetts 02118

Ralph Timperi, M.P.H.
State Laboratory Institute
Massachusetts Department of Public Health
305 South Street
Jamaica Plain, Massachusetts 02130

HISTORY OF HEROIN PREVALENCE ESTIMATION TECHNIQUES

Marc D. Brodsky

INTRODUCTION

Methods for estimating the number of heroin users have changed in response to changes in the nature of the opiates of addiction and their legality as well as to changes in types of available data. Greene (1974) and Rittenhouse (1977) have outlined some recent methods used to estimate the prevalence of heroin use. The part of the population affected by opiate addiction changed from the users of legal patent medicines to illegal opiates. The progression is traced by Austin (1979a and b) in a history of opiates in the United States from 1840 to 1930, and by Musto (1973). This paper reviews the techniques used to estimate the number of heroin users in the United States.

HEROIN PREVALENCE ESTIMATION TECHNIQUES PRIOR TO 1928

Several State- or community-based estimates of the number of opiate-dependent persons were made prior to 1928. Several sources of data were used other than interviewing addicts themselves. The techniques generally used were based on data provided by such sources as physicians, druggists, and treatment clinics. Terry and Pellens (1970) used these estimates to make national projections of the numbers of addicts for the prevailing period. They extended the following State- or community-based rates to the Nation:

Marshall (1878) surveyed physicians in 96 small towns in Michigan in 1877 and found 1,313 addicts. Terry and Pellens (1970) determined that if this survey in Michigan were representative of the United States, the national number of addicts would have been 251,936 in 1877.

Hull (1885) surveyed 123 druggists in Iowa in 1884, asking how many of their customers were opium addicts, and fourd 235 opium addicts. Again, Terry and Pellens (1970) found that if this survey in Iowa were representative of the United States, there would have been 182,215 addicts in the Nation in 1884.

Terry (1927) studied a treatment clinic population and the distribution of drug prescriptions in Jacksonville, Florida in 1913. Terry and Pellens (1970) stated that if the rate for Jacksonville, Florida based on the 541 opiate addicts found by Terry were to be extrapolated to the population of the entire country it would result in an estimate of 782,118 addicts.

Brown (1915) used State registration data in 1915 of 2,370 addicts in Tennessee to estimate the addict population. Terry and Pellens (1970) used Brown's Tennessee data to produce an estimate of 269,000 addicts in 1915.

The Secretary of the Treasury of the United States (1919) made an estimate of the number of addicts in treatment, based on a survey of 3,023 district, county, and municipal health officers. In this survey, there were 983 replies reporting a total of 73,150 addicts. Terry and Pellens (1970) found that if all physicians in the United States could be characterized by those who responded to this survey, the national estimate for the number of addicts in 1918 would have been 237,655.

Terry and Pellens' (1970) national projections for 1919 were based on the study by Hubbard, who made an estimate based on clinic data of 7,464 addicts in New York City. If this New York City data were representative of the country then, according to Terry and Pellens, the national estimate in 1919 would be 140,554 addicts.

Terry and Pellens' (1970) estimate for 1920 is based on the number of cases recorded at the Shreveport, Louisiana, treatment clinic over a period of 4 years. The Shreveport clinic director, W.P. Butler, recorded a count of 211 individual addicts in 1 year. The ratio of this number to the population of Caddo Parish, where the clinic was located, gives a rate of 0.25 percent. Terry and Pellens (1970) found that this rate, if representative of the United States, would imply 264,276 addicts in 1920.

METHODS DEVELOPED AFTER THE LATE 1960s

Between 1928 and the late 1960s, few estimates of the number of narcotic users were made. It was not until the number of heroin users increased in the late 1960s and early 1970s that estimates were again attempted.

Multiplier Methods for Estimating the Number of Heroin Addicts

A "multiplier" ratio relates some known estimate to the unkown desired estimate. Typically a multiplier is developed in a particular time period, then used without considering the potential for a change over time.

In 1969, Englander (Ball and Chambers 1970) established a multiplier based on the ratio of the number of addicts reported to the Bureau of Narcotics and Dangerous Drugs (BNDD) and those on the New York City narcotics register. This multiplier was then applied to the number of heroin addicts reported to the BNDD in 38 "heroin" states. A slightly different procedure was developed using BNDD and National Institute of Mental Health (NIMH) Lexington and Fort Worth treatment data for 12 southern States. The national total, 108,424 for 1967, was the sum of the estimate based on the 38 "heroin" States and the 12 southern States.

Friedman (1972) estimated the total number of habitual heroin users in the United States in 1970 to be 712,793. Friedman's method was based on an extrapolation of an estimate of the number of heroin users (230,000) in New York City. This number (230,000) is based on the number on the NYC narcotics register (150,000), corrected for the ratio of the number of reported narcotics deaths to cases known on the register (2.0), and several other factors: emigration (0.9), remission or death not noted on the register or duplicates on the register (0.9), and register underrepresentation of gainfully employed people with habits (1.05). The 712,793 is the sum of four parts: 1) the estimate for New York City; 2) the estimate for the total of all cities with a population of 500,000 to 5,000,000; 3) the estimate for the total for all cities with a population of 50,000 to 500,000, and 4) the estimate for the total of all cities with populations up to 50,000. Each of these parts is composed of three multiplied factors: the ratio of narcotics users in that area to New York by area size--1) 1.0, 2) 0.4, 3) 0.2, and 4) 0.1 respectively; the ratio of the total population in that area to the population of New York City-- 1) 1.0, 2) 3.02, 3) 5.25, 4) 16.47 respectively; and a heroin supply and criminal financial support factor of 1) 0.8, 2) 0.7, 3) 0.6, 4) 0.5 respectively.

Baden (1970) formulated the number of heroin addicts in a city by simply multiplying the number of heroin-related overdose deaths in a year by 100. Baden (1970) initially estimated that the overdose death records in New York City constituted approximately 1 percent of the known heroin addicts on the New York City narcotics register. The death rate among heroin addicts varies widely due to factors such as variations in the percentage of heroin and contaminants in retail packages, the user's physical condition, and the type and quantity of drugs used in combination with heroin. A national estimate of heroin addicts could not be made at that time because a national estimate of the number of heroin-related deaths was not available.

Survey Methods for Estimating Heroin Prevalence

The low rate of heroin addiction makes general household surveys impractical to use for estimating the number of heroin addicts (Harrell, Gfroerer, and Frank, this volume). Crider (this volume) has shown that the household population and the population of treated heroin addicts show similar trends in the year of first use of heroin. Data about the number of new heroin users might be useful in a model for calculation of the number of heroin addicts. Survey methodology relevant to heroin users has been advanced by Rittenhouse (1979), Lipton (1981, 1982, 1983), and Miller (1984).

Rittenhouse (1979) described an application of the randomized response technique. The respondent is asked to flip a coin or perform some other random choice for one of two questions. One of the questions concerns the respondent's heroin use. The other question concerns some nonthreatening behavior for which the response rate in the sampled population is known.

Fishburne (1979) developed the nominative technique, described also by Miller (this volume). This technique was established because it was believed that heroin use may be underreported in face-to-face household interviews. Application of this technique is done by asking respondents how many of their close friends who live in a household have used heroin. In order to remove duplicates from the count, information is also obtained regarding how many other close friends of a particular user also know about this person's heroin use. These data, in addition to the reported number of close friends, can generate an estimate of the number of heroin users in the sampled population.

Based on the 1982 household survey, it was estimated that 1.8 million people in the United States have ever used heroin in their lifetime (Miller et al. 1983). Sometimes prevalence estimates are reported as a percent of the population. For example, in a survey of high school seniors and followup cohorts, 1.2 percent of seniors were reported as having ever used heroin (Johnston 1982).

Issues of validity and population coverage are described by Johnston and O'Malley (this volume). Reliability and consistency in self-reports of drug use are described by O'Malley et al. (1983). The effect of truancy and high school dropouts on the estimation of the number of heroin users from high school surveys is discussed by Clayton and Voss (1982) and Kandel et al. (in press). The rate of heroin use is much higher among high school dropouts than among the sampled high school population (Robins and Murphy 1967). This fact must be considered when interpreting heroin use surveys in high school populations.

Miller (1984) has developed the Item Count Technique. Two random unique samples of the population are presented with a list of deviant behaviors. The difference between the two lists is that one list includes heroin use. The respondent simply states how many of the behaviors on the list he or she has done. The prevalence of heroin use is calculated from a function of the resultant probabilities.

Mathematical Models for Estimating Heroin Prevalence

The estimates of the number of heroin addicts made by Greenwood (1976), Person et al. (1977), Demaree et al. (1981), and Shreckengost (1983) cover the period from 1970 to 1982. These estimates are shown in table 1. Greenwood's (1971) indicator-dilution method extended the capture-recapture method for estimating the size of a hidden population. A "capture" is an occurrence of a name on the BNDD file of known heroin users. A "recapture" is a recurrence in a successive year. Greenwood's model describes the mathematical adjustment for the number of deaths and the increased probability of a recapture once a capture has taken place. The results are then multiplied by the proportion of addicted narcotic arrestees, assumed to be the same as the proportion of addicted narcotic treatment admissions, i.e., 77 percent. Greenwood's method results in an estimate of 546,000 heroin users in 1975. A multiple recapture model applied to repeated admissions to treatment for heroin addiction is described by Woodward (this volume). Heroin addiction recidivism and treatment readmission must be considered when applying such a model.

TABLE 1. *Estimates of number of heroin addicts United States 1969 through 1982*

Year	Greenwood	Person et al.	Demaree et al.	Shreckengost
1969	242,000			
1970	403,000	-	-	-
1971	430,000	-	-	-
1972	482,000	-	-	-
1973	464,000	573,000	-	434,000
1974	558,000	584,000	-	460,000
1975	546,000	540,000	-	475,000
1976	-	-	523,000	473,000
1977	-	-	495,000	475,000
1978	-	-	471,000	471,000
1979	-	-	420,000	463,000
1980	-	-	492,000	478,000
1981	-	-	-	488,000
1982	-	-	-	496,000

SOURCES: Greenwood 1976, Person et al. 1977, Demaree et al. 1981, Shreckengost 1983.

Levin et al. (1975) give a computer source code for a Systems Dynamics model of "The Heroin System" at the community level called the "Persistent Poppy Model." Their Persistent Poppy Model does not produce a national estimate but does provide considerable insight on the dynamic relations among the social parameters describing its operation. The nature of the heroin system depends upon characteristics of the local community. This may require that any national estimate based on models be composed of estimates that represent regions. The development of regional models may benefit from synthetic estimation techniques reviewed by Steinberg (1979).

Cooley et al. (1977) developed a Monte Carlo population simulation model for estimating the number of heroin users; the number of frequent heroin users; and the number in various states such as treatment, prison, or arrest. The prevalence estimates for 1976 were 1,796,000 ever used; 218,000 frequently use; 21,000 in treatment; 176,000 in prison; and 42,000 in arrest. This model helps tie together estimates from the household population survey and the estimates of frequent users from other techniques.

Several attempts have been made to extend the application of the indicator-dilution method to more than two samples in local estimates (Brazie 1978; Streeter 1981). Woodward and Doscher (1979) have used narcotic addict treatment data to make Standard Metropolitan Statistical Area (SMSA) estimates of the number of treated heroin addicts. Woodward and Ng (1979) have compared the result of their log-linear based capture-recapture estimates to the estimates made by Person et al. (1977) and found them to be highly correlated.

Person et al. (1977) developed a method based on the Heroin Problem Index (HPI) (Person et al. 1976). Five indicators were chosen to form a basis for an index of the heroin problem: 1) treatment program admissions per 100,000 population; 2) narcotic analgesic hospital emergency room episodes per 100,000 population; 3) heroin-related deaths per 100,000 population; 4) retail heroin price; 5) retail heroin purity. Rank order scores of 24 SMSAs were developed for each of these measures. A weight was assigned to each SMSA based on the sum of the rank order scores. This sum of rank order scores was called the HPI index value. The critical assumption at this point was that the HPI index value was a linear correlate of the rate of heroin use in each SMSA. The estimating method was as follows. An independent estimate of the heroin use prevalence rate for at least one SMSA with a large HPI index and at least one SMSA with a small HPI index was obtained. These two or more SMSAs were called anchor cities. A linear relationship was assumed to exist between the prevalence of heroin addiction in an SMSA and the value of the SMSA's HPI index. This linear relationship and the estimate of heroin prevalence for the two or more anchor cities determined the estimated value for the heroin prevalence for all the other SMSAs for which HPI index values were computed. The estimate for the United States was made based on a population extrapolation from the estimate for the 24 SMSAs. Person et al.'s method resulted in an estimate of 540,000 heroin addicts in 1975.

Systems Analysis Models, Dynamic Simulation

Shreckengost (1983) developed a Systems Dynamics model of the heroin supply/demand market based on international supply data and the number of susceptible individuals in the 14- to 34-year-old age group. The user population, as a portion of this susceptible population, was shown to change in response to changes in the availability of heroin in the United States. Shreckengost (1984) has developed a model that explains in part the relation of the numbers of small, medium, and large heroin users in terms of milligrams of pure heroin consumed per day. Shreckengost's method results in an estimate of 475,000 heroin addicts in 1975.

SUMMARY

Historical methods for estimating the number of heroin addicts were based on extrapolation from local surveys. More recent estimation methods have attempted to use data collected from special sources, such as records of narcotic arrests and treatment program admissions. In addition, national surveys of the household and high school populations have provided estimates of heroin use. Dynamic and stochastic models have been developed in attempts to relate data from the special populations and the surveys.

REFERENCES

Austin, G.A. Opiates, United States, 1840-1914. In: Perspectives on the History of Psychoactive Substance Use. National Institute on Drug Abuse Research Issues No. 24. DHHS Pub. No. (ADM) 79-810. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1979a. pp. 198-209.

Austin, G.A. Opiates, United States, 1914-1930. In: Perspectives on the History of Psychoactive Substance Use. National Institute on Drug Abuse Research Issues No. 24. DHHS Pub. No. (ADM) 79-810. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1979b. pp. 215-227.

Baden, M.M. Methadone-related deaths in New York City. Int J Addict, 5(3):489-498, 1970.

Ball, J. C., and Chambers, C. D. The Epidemiology of Opiate Addiction in the United States. Springfield Ill.: Charles C. Thomas, 1970. 337 pp.

Brazie, C. L. Study to Develop Methods of Estimating Populations of Urban Drug Abusers. Report, Contract No. DA 7AC002, Office of Drug Abuse Policy, Washington, D.C., 1978. 177 pp.

Brown, L.P. Enforcement of Tennessee anti-narcotic law. Am J Public Health, 5(4):323-333, 1915.

Clayton, R.R., and Voss, H.L. "Drug Use and Abuse Among Student Dropouts." Paper presented at the Technical Review on Drug Abuse and Dropouts, National Institute on Drug Abuse, Washington, D.C., 1982.

Cooley, P.; Rao, V.; Greenberg, S.; Schlenger, W.; and Shah, B. A Description and Application of a Computer Model for Forecasting Drug Abuse Patterns. Research Triangle Park, N.C.: Research Triangle Institute, 1978. 115 pp.

Demaree, R.G.; Hudiberg, R.A.; and Fletcher, B.W. Estimates of the Prevalence of Heroin Use in 24 Metropolitan Areas 1976-79. Fort Worth, Tx.: Texas Christian University Institute of Behavioral Research, 1981. 59 pp.
Doscher, M., and Woodward, J.A. Estimating the Size of Subpopulations of Heroin Users: Applications of Log-Linear Models to Capture Recapture Sampling. Los Angeles: University of California, 1979. 26 pp.
Fishburne, P. "Heroin Estimator Development: The Nominative Technique for Heroin Prevalence Estimation: Conclusions and Recommendations." Unpublished report submitted to the National Institute on Drug Abuse, 11/27/79, pursuant to Contract No. 271-78-3514. 7 pp.
Friedman, C.S. "The Use of Heroin in the United States in 1970." Unpublished internal report, Office of Planning and Program Evaluation, Office of the Secretary, U.S. Department of the Treasury, Washington, 1972. 70 pp.
Greene, M.H. Estimating the Prevalence of Heroin Use in the Community. Special Action Office for Drug Abuse Prevention, Executive Office of the President, Monograph Series A, No. 4. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1974. 20 pp.
Greenwood, J.A. Estimating the Number of Narcotic Addicts. Drug Control Division, Office of Scientific Support, Bureau of Narcotics and Dangerous Drugs, U.S. Department of Justice Document No. SCID-TR-3. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1971. 21 pp.
Greenwood, J.A. "Application of Greenwood's Capture/Recapture Technique to the U.S. Addict Register Data." U.S. Drug Enforcement Administration. Internal Memorandum dated 020176.
Hull, J.M. The opium habit. Biennial Report Iowa State Board of Health. 1885.
Johnston, L.D., Bachman, J.G., and O'Malley, P.M. Student Drug Use, Attitudes, and Beliefs: National Trends 1975-1982. Prepared for the National Institute on Drug Abuse by the University of Michigan Institute for Social Research. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1982. 134 pp.
Kandel, D.B.; Raevis, V.W.; and Kandel, P.I. Continuity in discontinuities: Adjustment in young adulthood of former school absentees and school dropouts. Youth and Society, in press.
Levin, G.; Roberts, E.V.; and Hirsch, G.B. The Persistent Poppy. Cambridge: Ballinger Publishing Co., 1975. 229 pp.
Lipton, D.S. Drug Use Among College Students in New York State. New York State Division of Substance Abuse Services. 1981. 49 pp.
Lipton, D.S. The Use of Psychoactive Prescription Drugs Among New York State's Household Population. New York State Division of Substance Abuse Services. 1982. 24 pp.
Lipton, D.S. Drug Use Among Tenants of Single Room Occupancy Hotels in New York City. New York State Division of Substance Abuse Services. 1983. 31 pp.
Marshall, O. The Opium Habit in Michigan. Annual Report Michigan State Board of Health, 6:61-73, 1878.
Miller, J.D. "A New Survey Technique for Studying Deviant Behavior." Ph. D. diss., George Washington University, 1984.

Miller, J.D.; Cisin, I.H.; Keaton, H.G.; Harrell, A.V.; Wertz, P.W.; Abelson, H.I.; and Fishburne, P.M. National Survey on Drug Abuse: Main Findings 1982. DHHS Pub. No. (ADM)83-1263. Washington, D.C.: Supt. of Docs., U.S. Gov.. Print. Off., 1983. 141 pp.

Musto, D.F. The American Disease: Origins of Narcotic Control. New Haven: Yale University Press, 1973. 354 pp.

O'Malley, P.M.; Bachman, J.G.; and Johnston, L.D. Reliability and consistency in self-reports of drug use. Int J Addict, 18(6):805-824, 1983.

Person, P.H.; Retka, R.L.; and Woodward, A.J. Toward a Heroin Problem Index: An Analytical Method for Drug Abuse Indicators. 1976. DHEW Pub. No. (ADM) 76-367. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1979. Stock No. 017-024-00522-3. 33 pp.

Person, P.H.; Retka, R.L.; and Woodward, A.J. A Method for Estimating Heroin Use Prevalence. DHEW Pub. No. (ADM) 77-439. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1977. 14 pp.

Rittenhouse, J.D. The Epidemiology of Heroin and Other Narcotics. National Institute on Drug Abuse Research Monograph 16. DHEW Pub. No. 017-024-00690-4. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1977. 249 pp.

Rittenhouse, J.D. Development Papers: Attempts to Improve the Measurement of Heroin Use in the National Survey Administrative Report, National Institute on Drug Abuse, Washington, D.C.: Supt. of Docs. U.S. Govt. Print. Off., 1979. 170 pp.

Robins, L.N., and Murphy, G.E. Drug use in a normal population of young Negro men. Am J Public Health, 57(9):1580-1596.

Schlenger, W.E., and Greenberg, S.W. Analysis of Intermetropolitan Correlates of Heroin Use. Report, Contract No. 271-75-1016, National Institute on Drug Abuse, Rockville, Md. 1978. 98 pp.

Shreckengost, R.C. Heroin: A New View. Washington, D.C.: U.S. Central Intelligence Agency, 1983. 25 pp.

Shreckengost, R.C. How Many Heroin Users Are There? Washington, D.C.: U.S. Central Intelligence Agency, 1984. 13 pp.

Steinberg, A., ed. Synthetic Estimates for Small Areas. National Institute on Drug Abuse Research Monograph No. 24. Washington, D.C.: Supt. of Docs., U.S. Gov. Print. Off., 1979. 282 pp.

Streeter, L.L. A Methodology for Prevalence Estimation and Computer-Generated Prevalence Maps. Report, Contract No. 271-80-4806, Drug Abuse Geoprocessing Methodology Project, National Institute on Drug Abuse, Rockville, Md., 1981. 141 pp.

Terry, C.E. 1913 Annual Report of the Jacksonville, Florida, Board of Health. In: The Report to the Committee on Drug Addictions. New York: Bureau of Social Hygiene, 1927.

Terry, C.E., and Pellens, M. The Opium Problem. New York: Bureau of Social Hygiene, 1928. Publication No. 115: Patterson Smith Reprint Series in Criminology, Law Enforcement, and Social Problems. Montclair, N.J.: 1970. 1,042 pp.

U.S. Treasury Department. Special Committee of Investigation. Traffic in Narcotic Drugs. Washington, D.C.: U.S. Govt. Print. Off., 1919.

Woodward, J.A., and Ng, L. "Structural Equation Models of Heroin Abuse Prevalence and Availability." Unpublished technical report written for the Drug Enforcement Administration in conjunction with 1979 DEA funding of the supplement to Public Health Service Grant DA 107060. University of California at Los Angeles, 1979. 32 pp.

Woodward, J.A., and Doscher, M.L. "Estimation of the Size of Different Subpopulations of Heroin Abusers." Unpublished technical report. University of California at Los Angeles, 1979. 14 pp.

AUTHOR

Marc D. Brodsky, M.S.
Division of Epidemiology and Statistical Analysis
National Institute on Drug Abuse
5600 Fishers Lane
Rockville, MD 20857

THE NOMINATIVE TECHNIQUE: A NEW METHOD OF ESTIMATING HEROIN PREVALENCE

Judith Droitcour Miller

The nominative technique is a relatively new method of indirect survey-based estimation that is being developed expressly for the purpose of estimating heroin prevalence in the general population. This new technique, which involves asking respondents to report on their close friend's heroin use, is essentially an attempt to reap the benefits of survey research, while at the same time avoiding some of the major problems of the self-report method. The primary purpose of the nominative technique is to minimize respondent denial of socially undesirable behavior. Another possible advantage is achieving coverage of "hard-to-reach" deviant population groups. The nominative question series has been inserted in the 1977, 1979, and 1982 National Surveys on Drug Abuse (Miller et al. 1983). The resulting nominative estimates of heroin prevalence are presented here and contrasted with corresponding self-report estimates.

BACKGROUND

As delineated in other papers in this volume, valid self-report data on serious forms of deviance and drug use are difficult to obtain via conventional forms of survey research. Of course, as in any survey, there are bound to be problems of respondent recall, incomplete knowledge concerning specific drugs used, etc., as well as some differences due to question wording. But these relatively routine difficulties represent far less serious sources of bias than two special problems faced by deviance researchers: First, the most deviant persons may not be captured by conventional methods of survey sampling. Second, persons who have engaged in serious forms of deviance, such as heroin use, have a very clear motive for distorting the facts of their experience in a survey interview.

Of these two problems, respondent denial has usually received the greater amount of attention, perhaps because each of us can identify, to some extent, with the situation of the deviant respondent. We know that if we were asked to disclose "undesirable facts" about ourselves, we would--consciously or unconsciously--anticipate some degree of negative consequence, perhaps merely the lowering of our self-image in the eyes of another person. An even stronger motive for denial would seem to exist whenever there is any possibility of more severe consequences, such as unemployment, job loss, or even prosecution or incarceration.

Cisin and Parry (1980) report a study of the validity of self-reported drug use, including heroin use. This work indicates that some heroin users do reveal the true facts of their experience, but many others do not--even if interviewers provide them with secret answer sheets, sealed envelopes, and sincere promises of confidentiality. Other empirical studies of response validity, as well as theories of social desirability and self-disclosure that are reviewed by Harrell (this volume) also suggest respondent denial of stigmatized behaviors.

Despite problems of denial of deviant behaviors such as heroin use, investigators have been reluctant to give up the possibility of obtaining valid survey data even in very sensitive question areas. As a result, over the past 20 years, a variety of novel survey-based estimation techniques have been devised with the purpose of encouraging deviant respondents to tell the truth. The first of these was the randomized response method, which dates from 1965 (Warner 1965; Folsom et al. 1973). Under the randomized response condition, a respondent might be shown <u>two</u> questions, each printed on a card. One card would display <u>the</u> sensitive research question, such as "Have you used heroin during the past month?" The other card would display an innocuous question, such as, "Were you born in April?" The respondent is instructed to answer only one of these questions and to determine <u>which</u> question by a secret randomizing device (such as flipping a coin in private). Then he (or she) merely says yes or no, and no one else, not even the interviewer, can ever know which of the two questions is being answered.

Randomized response is a form of self-reporting that avoids complete disclosure. Interviewers and researchers never know whether a particular respondent actually engaged in the deviant behavior that is being studied. Of course, once the results are combined for all respondents, it is possible to estimate the overall prevalence of the deviant behavior, based on prior knowledge of the prevalence of the innocuous behavior as well as the probability of selecting the deviant item. Unfortunately, there have been various problems reported by those who have implemented randomized response in the field. Apparently, some respondents do not comprehend <u>why</u> they are being asked to "play the game" and therefore may be suspicious and confused (Wiseman et al. 1975-76). Moreover, once the question of one's own deviant behavior has been raised, some innocent respondents apparently do not like being forced to remain ambiguous on that point. They would prefer to go on record with a statement that they have <u>not</u> engaged in it. Some of these people would say no regardless of the question drawn (Shimizu and Bonham 1978; Miller 1981). Another drawback of the randomized response method is a very high variance cost, i.e., the variance of a randomized response estimate is, in most cases, at least four times as high as the variance of a corresponding conventional estimate (Miller and Cisin 1980).

Because of the disadvantages of the randomized response approach, other methods of indirect self-reporting have been developed. These include the aggregated response method (Warner 1971), in which the respondent adds a random number to his (or her) quantitative answer. Another version of aggregated response features two subsamples; in one subsample respondents add responses to two (or more) questions, while respondents in the other are instructed to subtract responses (Boruch and Cecil 1979). The most recent version of indirect self-reporting, the "item-count/paired lists" technique (Miller 1983), is completely unobtrusive. Respondents in one subsample are shown a short list of behavior items (including the deviant behavior) and are asked <u>how many</u> of these categories apply to them; respondents in the other subsample are treated in exactly the same way, except that the deviant item is omitted from the list.

The nominative technique, described in the following section, is a rather different approach, for it involves reporting not on one's own behavior, but the behavior of <u>other persons</u> that one knows. This completely avoids the issue of whether or not the respondent has engaged in the behavior. The underlying premise is that respondents will be more truthful in reporting the deviance of anonymous others than they are in reporting their own deviant behavior. The nominative approach also has a unique potential for increasing coverage of deviant population groups.

THE NOMINATIVE METHOD: STATISTICAL LOGIC

The nominative technique is a variant of the multiplicity methods of survey research pioneered by Monroe Sirken, originally for the purpose of studying rare diseases and conditions (Sirken 1975). These multiplicity methods require the respondents to serve as informants reporting on the illness, behavior, or experience of (anonymous) other persons, such as their friends or relatives.

As applied to socially acceptable illnesses, the multiplicity approach asks about the respondent's own experience as well as that of close relatives, e.g., brothers and sisters. Knowing the number of brothers and sisters that a respondents has is the equivalent of knowing the number of siblings that can report each of the other's illnesses; this allows the investigator to "correct" for the duplication of reports of a single case, as occurs, for example, when two siblings both report that a third has a rare disease or condition.

In applying multiplicity methods to socially unacceptable or deviant behaviors, it is preferable to avoid requiring reports of the respondent's own behavior as well as that of close relatives. Instead, the focus in studies of deviance has been on the anonymous friends, or close friends, of the respondent. The version that was first applied to deviant behavior asked respondents, for example, "What proportion of your friends have used heroin?" In order to achieve a prevalence estimate, the investigator simply averaged the percentage of user-friends across all respondents. (Sudman et al. 1977). This approach assumes that deviants and nondeviants have equal numbers of friends--and that persons who know deviants have the same number of friends as persons who do not know deviants. It also assumes that the respondent knows whether each has engaged in the deviant behavior.

The newer version of this approach, which has been termed the "nominative technique," is a statistically defensible method of obtaining indirect estimates of deviant behavior. The technique was developed expressly for the study of heroin use, under contract to the National Institute on Drug Abuse. With this technique, the reference group is close friends.

The nominative technique is based upon the following proposition: If each member of the population reports the number (0, 1, 2, 3...) of close friends who have used heroin, then, with appropriate correction for duplication, it is possible to derive an accurate count of the number of heroin users in the population.

The two key items in the nominative question series are (in essence):

A. So far as you know, how many of your close friends have ever used heroin? Just count the ones that you know for sure have used it.

Then for each of the interviewee's heroin-using close friends:

B. How many of <u>this person's</u> other close friends (besides yourself) know that he (or she) has used heroin?

The information gathered in Question B allows an appropriate "weight" (or correction for duplication) to be attached to each report of a heroin-using friend. This weight is the inverse of the total number of persons in the population who are eligible to report that particular heroin user.

Specifically, the fractional weight that must be attached to the j^{th} interviewee's report of the i^{th} heroin user is:

$$1/(1+B_{ij})$$

This weight corrects for the fact some heroin users will be reported by two, three, four, or more close friends. To grasp this point, suppose for a moment that I am a heroin user in a population where a nominative census is being taken. Further, suppose that 10 of my close friends--also in this population--know that I have used heroin, and that each of them reports this in the interview. This would add up to a duplicated count of 10 heroin users instead of 1.

But now suppose that each of my friends tells the interviewer that nine others also know about my heroin use. Then the investigator would realize that each of the reports of my heroin use that would be included in the data would really be 1 of 10 reports of the same thing, and so each separate report should received a weight of 1/10. In other words, for each report of my heroin use, the investigator should count one-tenth of a heroin user. Then, when all 10 of these reports are counted together, they sum up to a total of 1 heroin user, which is what we started with. Taking one more example, suppose that only two people know about your heroin use. If each of these reports is counted as one-half of a heroin user, then the total of the two reports is simply 1 heroin user, which again is what we started with. And, of course, what can be done for each individual user can be done for all users in the population, so that the grand total of weighted counts of users would be the same as the total number of users in the population. The algebra of the nominative count in a population census is presented in appendix 1 of this paper. As described in a subsequent section and detailed in appendix 2, sample estimates are easily obtained.

THE NOMINATIVE TECHNIQUE: FIELD EXPERIENCE

Are people willing and able to provide answers to Question A and Question B with some degree of accuracy?

Prior to pretesting and field work, it was anticipated that Question A might be somewhat sensitive, that respondents might not want to talk about the heroin use of their close friends. Pretests revealed that people were quite willing, perhaps even eager, to talk about their friends' heroin use (Fishburne 1980). Moreover, when respondents were asked how they knew that their close friends had used heroin, they usually said that the friend himself (or herself) had told them. Sometimes, they said they had actually seen the person take heroin. In at least one case, the respondent said that he and his friend had "shot up" together. Roughly 10 percent of the total (nominative-form) sample in the 1977, 1979, and 1982 National Surveys on Drug Abuse reported that one or more of their close friends had used heroin.

Quantitative analyses of the National Survey data revealed that drug users were much more likely than nonusers to report having close friend(s) who had used heroin (Fishburne 1980). Thus, there have been several indications of valid responses to Question A.

Question B (How many of a particular heroin user's other close friends also know about his/her heroin use) is much more difficult. In each of the three surveys, 15 percent to 20 percent of the respondents who had a close-friend user could not estimate the number of that person's other close friends who also know. For this group of respondents, imputation procedures have been employed in data analysis. Specifically, based on the notion that all users known by a single respondent probably know each other, the rule for cases where Question B is not answered is to set the answer to Question B equal to the answer to Question A minus one. (The user himself/herself is subtracted out, since the counting rules for the nominative estimate do not include self-reports.) Applying this conservative rule ought not distort the data, but the fact that so many persons were unable to provide an estimate of B casts doubt on the accuracy of the responses that were given.

THE PRACTICAL VERSION OF THE NOMINATIVE TECHNIQUE

The foregoing discussion has emphasized that each heroin user in the population may be known to several interviewees. The other side of the coin is that any one respondent may report several heroin users. From a practical point of view, it can be awkward to ask Question B separately for each heroin user that a particular respondent knows. Of course, many respondents don't know any users at all and many others would know only one (Miller 1983). But for those who do know two or more heroin users, Question B must be asked separately for each user. For example, if a respondent tells the interviewee that he (she) knows five heroin users, then the interviewer must ask first for the first person, how many of his other close friends also about his heroin use; then, for the second person, how many of his friends know; and so forth for all five users. Moreover, it is necessary to ask about the age, sex, and recency of heroin use for each heroin user, so that estimates can be derived for age-sex subgroups and for current use. Obviously, this can become burdensome, especially if imbedded in an already lengthy interview.

For practical purposes, an optional version of the nominative question series was developed for use in the National Surveys on Drug Abuse. This limits the number of nominees that are followed up with Question B and the questions on sex, age, and recency of use. Specifically, in the National Survey, only one nominee (per respondent) is referenced in the followup questions. Appendix 3 of this paper presents copies of the exact questions and procedures that were used in the 1982 National Survey. Briefly, whenever a respondent reported knowing two or more heroin users, one of these persons was randomly selected in the following way. Interviewers gave respondents a folded card, so they could list the initials of their close friends who had used heroin--and number these persons. The interviewer then consulted a random number table and told the respondent that the rest of the questions would be about "Person Number___." Very little difficulty has been encountered in implementing these procedures in the field. In the first (1977) National Survey to include the nominative question series, a few respondents objected to writing their friends' initials on the card; in subsequent surveys, such persons were urged to number their friends "in their heads."

Statistically, the fact that a single nominee is selected from all those mentioned in Question A means that this single user--this "random one"--must stand for the total number of users the respondent knows. For example, if a respondent knows three heroin users ($A_j = 3$) and we randomly select one of them, then that one must stand for all three of them. This is accomplished in data analysis simply by "weighting up" the "random one." In other words, for the j^{th} interviewee, the random one must be counted A_j times. This procedure is analogous to the conventional practice of randomly selecting one adult respondent per household and then weighting up that lone respondent to stand for the entire group of adult residents in that household.

Thus, in the practical version of the nominative technique, each interviewee or respondent has either a "0" score for no users known or a "1" score for a report of a particular heroin user known. This particular user is either the only user that the respondent knows or is the "random one" selected from all users known to that respondent. This particular user, who has a number of characteristics such as sex, age, and recency of use, bears two data weights:

A_j which "weights up" the j^{th} respondent's report of this heroin user to stand for the total number of users that he or she has mentioned.

$1/(1+B_j)$ which is the correction for duplicate reports of the same user described above.

The combined weight factor for each score of "1" (i.e., for each respondent who knows a heroin user) is:

$$A_j \quad (1/(1+B_j))$$

(Note: The i subscript has been dropped, since in this version there is only one user referenced per respondent.)

Using this combined weight factor--and summing reports across all respondents in a complete census of a given population--yields an estimate of the total number of heroin users in that population.

OBTAINING NOMINATIVE PREVALENCE ESTIMATES FOR SUBGROUPS AND SAMPLES

Once the total nominative count of heroin users for a population is obtained, dividing that figure by the total population size produces a prevalence rate or percentage. For a subgroup, such as females, the analyst would simply sum those reports in which the random heroin user is a female--across both male and female respondents (for it is the characteristics of the nominated users that we are interested in, not the characteristics of the respondents who are nominating them.) Once a nominative count of all female users in the population has been obtained, this number may be divided by the total number of females in the population to produce a prevalence rate.

To this point, the statistical discussion has assumed a complete census of the population. Actually, in most research (as, of course, in the National Surveys in which the nominative question series has been inserted), sample data are used to estimate population prevalence figures. When using nominative data from a sample of respondents, the analyst would simply take the sample count of nominated heroin users, weighting each report with the combined weight factor shown on the previous page. Then, to obtain a population projection, an inflation factor would be used to raise the nominative sample count to the population level. The inflation factor is simply N/n where N refers to the population size and n is the sample size; this is the inverse of the sampling fraction. This procedure is shown in appendix 2 of this paper.

For a prevalence estimate based on sample data, the analyst simply divides the total nominative count for the sample by the total sample size (appendix 2). When deriving a <u>subgroup</u> prevalence estimate from nominative sample data, one would simply count all (weighted) reports in which the nominee is a member of the relevant subgroup, e.g., is a female; then, this count would be divided by the number of <u>respondents</u> in the sample who are members of that subgroup, e.g., the number of female respondents (appendix 2). Throughout, the usual data weights used for conventional analyses of the sample data would be applied.

The variance of a sample-based nominative estimate depends upon variability in the nominative weight, i.e., $A_i(1/(1+B_i))$. Thus, it is not possible to express the variance associated with the nominative estimator as a function of the size of the variance of a corresponding direct estimator. Empirical variance estimates have been calculated via random subsampling, using data from the pilot test of the nominative technique (in application to heroin estimation) in the 1977 National Survey on Drug Abuse. Cisin (1980) has reported these nominative variances in comparison to comparable direct-report variances. The preliminary results indicate that, as applied to the estimation of heroin use, the nominative technique does not, in general, carry a high variance cost. However, given the practical version of the nominative approach that has been used in the National Survey (followup questions for only one nominee per respondent), even a few very high A_i values could raise the variance substantially. In order to discourage overreporting of heroin-using close friends, the nominative question series used in the 1979 and 1982 surveys features the introductory question: "How many close friends do you have?" This tends to limit the number of close-friend heroin-users that will be subsequently reported. In general, the nominative approach carries much lower variance costs than the methods of indirect self-reporting described in an earlier section of this paper.

NOMINATIVE HEROIN ESTIMATES FROM THE 1982 NATIONAL SURVEY AND TWO PREVIOUS NATIONAL SURVEYS

Table 1 shows nominative and self-report estimates of lifetime heroin prevalence from the 1982 National Survey on Drug Abuse. Clearly, the nominative estimates are higher for young adults than for other age groups, and within each age group, the estimates are higher for males than for females. These basic patterns follow expected distributions, based on a variety of data sources. Moreover, table 1 indicates that for each sex-age group, the nominative estimates are higher than the corresponding self-report estimates. Notably, for young males, the nominative estimate indicates that 9.4 percent have at least tried heroin at some point in their lives, whereas the self-report estimate is only 1.4 percent. These are the kinds of patterns that were expected if people were, in fact, more likely to tell the truth about their friends' heroin use than about their own heroin use.

TABLE 1. *Nominative and self-report estimates from the 1982 National Survey on Drug Abuse: Lifetime prevalence of heroin use by sex and age*

Population by (Age) and Sex	Estimate of Ever Used Heroin by Percent	
	Nominative	Self-report
Youth (12-17)		
Males	4.5	0.6
Females	1.8	*
Young Adults (18-25)		
Males	9.4	1.4
Females	5.6	0.9
Older Adults (26+)		
Males	2.7	2.0
Females	0.9	*

*Less than one-half of one percent; estimate not shown.

Note: Self-reports of heroin use are based on the entire 1982 survey sample of 1,581 youth; 1,283 young adults; and 2,750 older adults. Nominative estimates are based on the answers of the 2,852 respondents who were assigned the N-form questionnaire, which included the nominative question series; 211 of these respondents reported having one or more close friends who had used heroin. Of these 211 reports, 47 referred to a youth nominee, 85 to a young adult nominee, and 79 to an older adult nominee.

Table 2 provides nominative estimates of past-year and past-month heroin use for three age groups. Self-report data are not shown, because estimates were less than one-half of one percent in all age categories. Thus, the nominative technique produces higher estimates for current use than does the conventional self-report method. Population projections for current heroin users were not derived from self-reports in the 1982 National Survey, because of the low prevalence of self-reported current use. Population projections based on the nominative technique are, however, shown in table 2. The total population estimate for past-year users is nearly two million.

TABLE 2. *Nominative estimates from the 1982 National Survey on Drug Abuse: Past-year and past-month prevalence of heroin use and population projections for three age groups*

	Percent Used Heroin In: Past Year	Past Month	Number Used Heroin In Past Year
Youth 12-17	2.0	1.0	450,000
Young adults 18-25	2.6	1.0	840,000
Older adults 26+	0.5	*	590,000
Total number used			1,880,000

*Less than one-half of one percent; estimate not shown.

Note: Number of respondents providing self-report and nominative data is given in table 1 on the previous page. The number of nominative reports referencing a past-year heroin user was 29 for youth nominees, 43 for young adult nominees, and 21 for older-adult nominees; the corresponding figures for past-month use are 18, 19, and 9, respectively.

Table 3, on the following page, shows trends in nominative estimates of the lifetime prevalence of heroin use, from the 1977, 1979, and 1982 National Surveys on Drug Abuse. Self-report trends are also shown for purposes of comparison.

Perhaps the most striking feature of table 3 is that, consistently, across all 3 survey years, nominative estimates have been higher than corresponding self-report estimates.

Overall, the nominative data (part A of table 3) suggest a downward trend in the percentage of the population that has ever tried heroin. The downward trend apparently began with a slight decrease between 1977 and 1979 and then continued with a somewhat more marked decrease between the 1979 and 1982 studies. The apparent decrease in older adult's heroin use, as measured by the nominative method, should be viewed with caution. Based on patterns of illicit drug use in various age groups during recent years, the former young adults who recently moved into the older adult category are more likely to have tried illicit drugs than are elderly persons who "drop out" of the population because of death or institutionalization in nursing homes. The only alternative explanation would be untimely deaths or institutionalization of persons who tried heroin.

TABLE 3. *Nominative and self-report trends, based on the 1977, 1979, and 1982 National Surveys on Drug Abuse: Lifetime prevalence of heroin use in three age groups*

A. Nominative Trends: Percent Ever Used Heroin

	1977	1979	1982
Youth (12-17)	8.3	7.5	3.1
Young Adults (18-25)	10.7	9.1	7.3
Older Adults (26+)	4.0	3.6	1.7

B. Self-report Trends: Percent Ever Used Heroin

	1977	1979	1982
Youth (12-17)	1.1	0.5	*
Young Adults (18-25)	3.6	3.5	1.2
Older Adults (26+)	0.8	1.0	1.1

*Less than one-half of one percent; estimate not shown.

Note: For numbers of respondents on which self-report and nominative estimates from the 1977 and 1979 surveys are based, see Miller (1983).

Part B of table 3 also indicates that self-reports of heroin use based on the same three surveys also suggest a downward trend in heroin prevalence among young persons--but not among older adults. For reasons just stated, the self-report for older adults seems more reasonable than the nominative trends for this age group. The possible anomaly in older-adult ever-use trends, based on the nominative method, serves as a reminder that no strict tests of validation have been performed for the nominative approach.

SUMMARY

Over the years, nominative estimates of heroin prevalence have been consistently higher than self-reports of heroin use. During this time, nominative data have generally followed mainstream patterns of drug use: nominative estimates for young adults and for males are higher than nominative estimates for older persons, youth, and females; moreover, the recent downward trends in drug use have been replicated by the nominative heroin data. Thus, the overall picture presented by the nominative data--similar patterns but higher levels of prevalence--seems to support the validity of the new approach.

Nevertheless, considerable caution should be exercised in interpreting nominative data. This is chiefly because a substantial minority of nominators cannot report the number of other close friends of the heroin user who also "know." While missing data has been handled by a conservative imputation rule, the fact that so many persons are unable to provide an answer to this key question casts doubt on the accuracy of the answers that were given. In fact, the nominative approach might tend to produce overestimates, because of the potential for undercounts of the numbers of others who "know."

Additional tests of validity should be performed, such as application of the nominative approach to nonsensitive behaviors or minimally sensitive behaviors, such as marijuana use or perhaps cocaine use. Certainly, the overall validity of the nominative heroin data would be supported if in future surveys new nominative heroin estimates for relatively unstigmatized forms of drug use proved to be similar to self-reported levels of use, thus pointing to the unique difference in estimates that might be observed for heroin.

Finally, in interpreting the heroin estimates presented here, it should be remembered that both the nominative and self-report estimates refer to heroin use in the household population of the United States. Thus, many heroin addicts and other users who reside in various unconventional living arrangements would not be included in the counts presented here. Among the excluded groups are transients residing in rooming houses or "crashing" in the home of one "friend" after another or who are incarcerated in jails or confined to residential drug treatment centers. This is a caution for interpreting the estimates presented in this paper, not a criticism of the nominative technique itself. In fact, since many persons in the household population undoubtedly have friends who live in nonconventional residences, the nominative question series could be revised to include respondent reports of heroin users who presently reside in group quarters or who lack a fixed address.

To begin to reap the potential benefits of the nominative approach, further research is needed. First, investigations into the validity of nominative estimates are necessary, if these data are to be interpreted with confidence. Second, preliminary studies of revised question wording that would extend the reference population beyond the household population would greatly enhance the utility and unique advantages of this approach.

REFERENCES

Boruch, R.F., and Cecil, J.S. Method for Assuring Privacy and Confidentiality of Social Research Data. Philadelphia: University of Pennsylvania Press, 1979. 312 pp.

Cisin, I.H. The variance of prevalence estimates. In: Rittenhouse, J.D., ed. Developmental Papers: Attempts to Improve the Measurement of Heroin Use. Unpublished report. Rockville, Md.: National Institute on Drug Abuse, Rockville, Md., 1980. pp. 65-72.

Cisin, I.H., and Parry, H.J. Sensitivity of survey techniques in measuring illicit drug use. In: Rittenhouse, J.D., ed. Developmental Papers: Attempts to Improve the Measurement of Heroin Use. Unpublished report. Rockville, Md.: National Institute on Drug Abuse, 1980. pp. 3-46.

Fishburne, Patricia M. Survey techniques for studying threatening topics: A case study on the use of heroin. Doctoral Dissertation, New York University, 1980.

Folsom, R.E.; Greenberg, B.G.; Horvitz, D.G.; and Abernathy, J.R. The two alternate questions randomized response model for human surveys. J Am Stat Assoc, 68:525-530, 1973.

Miller, J.D. Complexities of the randomized response solution. (Comment on Tracy and Fox, ASR, April, 1981.) Am Socil Rev, 46:928-930, 1981.

Miller, J.D. A new survey technique for studying deviant behavior. Doctoral Dissertation, The George Washington University, 1984.

Miller, J.D. The nominative technique: Method and heroin estimates from the 1982 National Survey on Drug Abuse. Unpublished report. National Institute on Drug Abuse, Rockville, Md., October, 1983.

Miller, J.D., and Cisin, I.H. Measuring the hidden prevalence of heroin use. In: Rittenhouse, J.D., ed. Developmental Papers: Attempts to Improve the Measurement of Heroin Use. Unpublished report. National Institute on Drug Abuse, Rockville, Md., 1980. pp. 74-170.

Miller, J.D.; Cisin, I.H.; Gardner-Keaton, H.; Wirtz, P.W.; Abelson, H.I.; and Fishburne, P.M. National Survey on Drug Abuse: Main Findings 1982. DHHS Pub. No. (ADM) 83-1263. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1983. 141 pp.

Shimizu, I.M., and Bonham, G.S. Randomized response technique in a national survey. J Am Stat Assoc, 73:35-39, 1978.

Sirken, M.G. Network surveys of rare and sensitive conditions. In: Advances in Health Survey Research Methods. National Center on Health Statistics Research Proceedings Series, 1975. pp. 31-32.

Sudman, S.; Blair, E.; Bradburn, N.M.; and Stocking, C. Estimates of threatening behavior based on reports of friends. Public Opin Q, 41:261-264, 1977.

Warner, S.L. Randomized response: A survey technique for eliminating evasive answer bias. J Am Stat Assoc, 60:63-69, 1965.

Warner, S.L. The linear randomized response model. J Am Stat Assoc, 60:63-69, 1971.

Wiseman, F.; Moriarty, M.; and Schafer, M. Estimating public opinion with the randomized response mode. Public Opin Q, 39:507-513, 1975-76.

AUTHOR

Judith Droitcour Miller, Ph.D.
Social Research Group
The George Washington University
2000 L Street Suite 305
Washington, D.C. 20036

APPENDIX 1

The statistical logic of the nominative count is presented below. The notation to be used is as follows:

X = a sensitive attribute, such as heroin use; $X_i = 1$ if the i^{th} person in the population possesses the attribute (e.g., has used heroin); $X_i = 0$ if the i^{th} person does not possess the attribute.

T_X = the total number of persons in the population who possess the attribute; i.e., $T_X = \Sigma X_i$.

F = an attribute involving an interviewee's ability to report that another person possesses attribute X. F_{ij} refers to the j^{th} interviewee's report of the i^{th} person's heroin use. Thus, $F_{ij} = 0$ if the j^{th} person does not report that the i^{th} person is a heroin user. $F_{ij} = 1$ if the j^{th} person does report that the i^{th} person is a user. (Note: $i \neq j$; i.e., no self-reports are included.)

$\sum_{i \neq j}^{N} \sum^{N} F_{ij}$ = the total number of reports that persons possess attribute X. (Obviously, if uncorrected, this would be an overcount--unless each user is known to only one close friend.)

$\sum_{j=1}^{N} F_{ij}$ = the number of times the i^{th} person is reported as a heroin user or the number of interviewees who report this particular heroin user.

$\sum_{i=1}^{N} F_{ij}$ = the number of "nominees" (possessors of attribute X) reported by the j^{th} person. In other words, this term represents a sum across heroin-user nominees (i's) for the j^{th} nominator.

The total number of heroin users in the population can be expressed as follows:

$$T_X = \sum_{i=1}^{N} X_i = \sum_{i=1}^{N} \left[X_i \left(\sum_{j=1}^{N} F_{ij} \right) / \left(\sum_{j=1}^{N} F_{ij} \right) \right]$$

Since:

$$\sum_{i=1}^{N} X_i \left(\sum_{j=1}^{N} F_{ij} \right) = \sum_{i \neq j}^{N} \sum^{N} F_{ij}$$

That is, each user in the population, times the number of others who report him, summed across all users, totals to the total number of reports of heroin users.

Therefore:

$$T_X = \sum_{i \neq j}^{N} \sum^{N} (F_{ij} / \sum_{j=1}^{N} F_{ij}) \qquad \text{Letting } 0/0 = 0 \text{ if } \sum_{j=1}^{N} F_{ij} = 0$$

Thus, in the population, the total number of persons possessing attribute X (i.e., T_X) is equal to: the sum of the nominations (reports of other persons possessing attribute X), where each individual report of a user is divided by (weight by the inverse of) the number of persons reporting that particular user. In other words, each report of a user is weighted by the inverse of the number of persons reporting that particular heroin user.

APPENDIX 2

In a survey in which a sample of size n is drawn from a population of size N, the sampling fraction is n/N. Any direct count of persons in the sample who possess a certain characteristic, i.e.:

$\sum_{j=1}^{n} C_j$ where $C_j = 1$ if the jth person possesses the characteristic and otherwise is equal to 0,

can be raised or "projected" to a population value via applying the inverse of the sampling fraction. That is:

$$\text{EST } (T_C) = N/n\left(\sum_{j=1}^{n} C_j\right)$$

The same reasoning holds for the nominative count. Letting Z_j = the weighted nominative score of the j^{th} respondent:

$$\text{EST } (T_X) = N/n\left(\sum_{j=1}^{n} Z_j\right)$$

In order to obtain a prevalence estimate based on a nominative count, one could divide the population projection by the total population size, i.e.:

$$\text{EST } (\mu_X) = \frac{\text{EST } (T_X)}{N}$$

However, this expression can be reduced as follows:

$$\text{EST } (\mu_X) = \frac{N/n\left(\sum_{j=1}^{n} Z_j\right)}{N} = (1/n)\left(\sum_{j=1}^{n} Z_j\right) = \overline{Z}_j$$

Thus, the mean nominative score for the entire sample can be used as an unbiased estimate of heroin prevalence in the population from which the sample was drawn.

CLOSE FRIENDS (FORM N ONLY)

39. Now, we would like you to think about people you know who live in regular households. Please do not include those people who live in a college dormitory, on a military base, jail, in a drug rehabilitation center, or have no definite address. Ready?

Most of us know many people. But, usually only some of these, if any, are people that we consider to be close friends. About how many close friends would you say that you have? Remember, we are only interested in those close friends who live in regular households.

________ NUMBER OF CLOSE FRIENDS LIVING IN REGULAR HOUSEHOLDS *(ACCEPT ONLY A NUMBER)* *342-343*

0 NO CLOSE FRIENDS LIVING IN REGULAR HOUSEHOLDS *SKIP TO PAGE 34*

40. This next question is about your ________ *(INSERT NUMBER FROM Q. 39)* close friends who live in regular households. Keep the names of these people to yourself. We want to know about them, but we do not want to know who they are.

About how many of these close friends can you say for sure have ever used heroin? We want to know about them, but we do not want to know who they are, because we are going to ask you about their drug use.

________ NUMBER OF CLOSE FRIENDS LIVING IN REGULAR HOUSEHOLDS WHO EVER USED HEROIN *(ACCEPT ONLY A NUMBER)* *344-345*

0 NO CLOSE FRIENDS LIVING IN REGULAR HOUSEHOLDS WHO EVER USED HEROIN -- *SKIP TO PAGE 34*

INTERVIEWER: IF RESPONDENT HAS ONLY ONE CLOSE FRIEND WHO HAS USED HEROIN, GO TO Q. 41, TOP OF PAGE 33. IF MORE THAN ONE FRIEND, GO TO TOP OF NEXT PAGE.

INTERVIEWER: IF MORE ONE THAN ONE CLOSE FRIEND WHO USED HEROIN, READ THE FOLLOWING:

GIVE RESPONDENT A SMALL BLANK WHITE CARD

On the card I gave you, I would like you to list the initials of your _______ *(INSERT FROM Q 40)* close friends who live in regular households who you know for sure have ever used heroin. No one but you will ever see these initials. *(WAIT UNTIL RESPONDENT MAKES LIST. IF RESPONDENT REFUSES TO USE CARD, HE/SHE MAY DO THIS PART IN HIS/HER HEAD.)*

Now, please number the people on your list. Put the number "one" next to the initials of the first person on your list. Then put the number "two" next to the initials of the second person on your list, and so on until everyone on your list has a different number. *(WAIT UNTIL RESPONDENT FINISHES NUMBERING.)*

I only want to ask you about one of the persons on your list. *(INTERVIEWER: USE TABLE BELOW TO SELECT CORRECT INDIVIDUAL.)*

INTERVIEWER: CIRCLE NUMBER OF PERSON YOU ARE GOING TO ASK ABOUT. THAT IS THE ONLY PERSON TO ASK ABOUT. NO SUBSTITUTES.

IF THE NUMBER OF CLOSE FRIENDS IN Q. 40 IS:	ASK ABOUT PERSON NUMBER
2	2
3	3
4	2
5	1
6+	6

Please draw a circle around the initials of the person number ____ *(INSERT FROM TABLE)*; the remaining questions will be about this person.

Rotation 6 of 6

41. Is this person male or female?
  1 MALE
  2 FEMALE *346*

42. How old is this person now? Is he/she 12-17 years old, 18-25 years old, 26-34 years old, or more than 34 years old? 8 NOT SURE
  1 12-17 YEARS OLD
  2 18-25 YEARS OLD
  3 26-34 YEARS OLD *347*
  4 35+ YEARS OLD

43. As far as you know, how long ago was the first time this person tried heroin?
  1 WITHIN THE PAST MONTH
  2 WITHIN THE PAST YEAR *348*
  3 MORE THAN A YEAR AGO
  8 NOT SURE

44. As far as you know, when was the most recent time this person used heroin?
  1 WITHIN THE PAST MONTH
  2 WITHIN THE PAST YEAR *349*
  3 MORE THAN A YEAR AGO
  8 NOT SURE

45. There are many different ways of knowing that another person has used heroin. Please tell me how you know for sure that this person has used heroin. *(WRITE EXACTLY WHAT RESPONDENT SAYS. IF RESPONDENT SAYS "SOMEONE ELSE TOLD ME" OR "EVERYBODY KNOWS," RECORD VERBATIM, THEN PROBE:* How do they know?)

350

46. Now, we would like you to think about this person's other close friends, besides yourself.

As far as you know, how many of this person's other close friends, besides yourself, know for sure that this person has ever used heroin? Remember, we are only interested in his/her close friends who live in regular households. *(IF RESPONDENT FINDS QUESTION HARD TO ANSWER OR SAYS "ALL" OR "MANY OF HIS/HER CLOSE FRIENDS," SAY: We need to have a number; please give us your best estimate.)*

________ NUMBER OF CLOSE FRIENDS LIVING IN REGULAR HOUSEHOLDS WHO KNOW *352-353*

0 NO OTHER CLOSE FRIENDS LIVING IN REGULAR HOUSEHOLDS WHO KNOW

* COULD NOT MAKE AN ESTIMATE

HEROIN INCIDENCE: A TREND COMPARISON BETWEEN NATIONAL HOUSEHOLD SURVEY DATA AND INDICATOR DATA

Raquel A. Crider

INTRODUCTION

A commonly held belief is that respondents in a face-to-face survey will underreport the extent of their involvement in deviant behavior (Sirkin 1975, Fishburn 1980, Miller this volume). To test this assumption, a time series of self-reported year of first use of heroin was compared to a time series of heroin indicator data, such as heroin-related emergencies and deaths, treatment admissions, hepatitis B cases, etc. Although there is a general belief that self-reported heroin use would be underreported, no test of that assumption has been conducted comparing the time series trend of heroin incidence based on survey data to the time series trends of heroin indicators.

PROCEDURE

Seven drug indicators are compared: 1) the number of heroin initiates based on the face-to-face National Household Survey on Drug Abuse (National Survey); 2) the number of heroin initiates voluntarily entering a panel of consistently reporting federally funded treatment programs (Client Oriented Data Acquisition Process, CODAP) for the first time by the year of first use; 3) the residual number of hepatitis B cases per year; 4) the percent of high school seniors ever using heroin (Johnston et al. 1982); 5) the number of heroin-related emergency room visits reported to the Drug Abuse Warning Network (DAWN); 6) the number of heroin-related deaths reported to DAWN; and 7) the average street level heroin purity (DEA 1984).

The analytic technique used in this paper produces numbers to examine the relationship between the National Household Survey trends and other indicators of new heroin users. The generated numbers are not to be construed as actual numbers of new heroin users in any particular year.

The household survey year-of-first-use data, the number of hepatitis B cases, and the number of heroin users in high school (recent initiates based on their age) are considered incidence indicators. The number of emergency room visits and heroin-related deaths are considered prevalence indicators. Heroin purity is not considered to be either an incidence or prevalence indicator, but is presented as a correlate of heroin incidence and prevalence.

Data from the 1977, 1979, and 1982 National Household Surveys on Drug Abuse were pooled for the analysis presented in this report. The total pooled sample consisted of 17,000 interviews. A total of 274 heroin users were found (1.6 percent of the pooled sample). The data from the national surveys are presented in 2-year moving averages for purposes of data smoothing and all discussion is based on these averages. Additional data smoothing is based on spline interpolation (Integrated Software Systems Corporation 1983).

While the purpose of this paper is to show a relationship between the number of self-reported heroin users over time and the number of hepatitis B cases, the number of treated initiates, and other heroin indicators, the purpose is not to make a statement about the number of heroin initiates in any particular year. With 17,000 interviews upon which to base the trend data and a total of 274 heroin users, it is thought that the trends over time can be interpreted and a correspondence between epidemic periods in the various indicators can be observed.

The epidemic periods discussed are those reported in the Heroin Work Group Report (NIDA 1984), i.e., the 1968-72 period (the early 1970s epidemic), the 1974-76 period (the mid-1970s epidemic), and the increase starting in approximately 1979-80 (the recent epidemic). For purposes of this report, an epidemic period is defined as those years in which several heroin indicators simultaneously increased and remained at elevated levels for 2 or more years.

Household Survey-Based Incidence Curve

Each of the three national household surveys contain items relating to the year of first heroin use and the age of first heroin use. The "age of first use" (AFU) item was used rather than the "year of first use" (YFU) item because it is believed that an age at which an event occurred is easier to remember than the year in which it occurred.

The AFU was converted to YFU by combining AFU with the year of the survey (YOS) and the age at the survey (AAS) using the equation: YFU=YOS-AAS+AFU.

Since the data were pooled over three surveys, the data were first weighted based on the survey sampling plan. Then each observation was weighted in accordance with the population represented. The population projections for 1977, 1979, and 1982 were then averaged. This pooling procedure takes into account the aging of the population from one survey to another and the differing sample sizes in each of the surveys, but does not attempt to minimize the variance estimates since sample variance estimates were approximately equal.

For the years of first use in 1978 through 1982, weighting factors were needed, since the 1977 survey could only include respondents who began heroin use prior to 1978, while the 1979 and 1982 surveys could only include respondents for whom heroin initiation preceded or occurred during those survey years. Thus, weighting factors of 1.5 and 3.0 were used for recent initiates in the 1979 and 1982 surveys respectively to adjust respondent rates to the 1977 base. The factor of 1.5 serves to weight two surveys as if they were three surveys for recent initiates. The factor 3.0 weights one survey as if it were three surveys for initiates starting after 1979.

Two-year moving averages were used throughout the period of analysis, i.e., for 1965 through 1980. The greatest variance about the 2-year moving mean occurred for years prior to 1973. The variance is thought to be due primarily to the difficulty of remembering the age of first use for an event occurring several years prior to the surveys.

Treatment-Based Incidence Curve

The year-of-first-use distributions from treatment admissions to federally-funded treatment programs are available through CODAP (1977-81 data were used in this report). The number of new users per year entering treatment for the first time is obtained from the year-of-first-use question on the treatment admission form. These data, when tabulated by year of first use, produce an incidence curve. Because recent initiates will wait for some time before entering treatment, a correction to the treatment data is needed. This correction procedure is based on the distribution of the lag time between first use and treatment for 1977 admissions (NIDA 1984b).

A recurring question is "What proportion of all heroin users enter treatment?" To address this question, a regression analysis was applied between the number of heroin initiates in households (2-year moving averages) and the 2-year moving average of the number of heroin initiates in treatment by year of first use, fitting the model:

$$\begin{array}{l} \text{Number of initiates in} \\ \text{households starting in} \\ \text{a particular year} \end{array} = B0 + B1 \times \begin{array}{l} \text{number of initiates} \\ \text{in treatment starting in} \\ \text{a particular year} \end{array}$$

where B0 is the number of initiates in households when the number of treatment initiates is zero. The B1 is the slope of the regression line. This slope is the factor by which the number of treated initiates would be multiplied to obtain the number of household initiates.

Hepatitis-Based Incidence Curve

The number of hepatitis B cases for 1975-83 was obtained from the Centers for Disease Control. Hepatitis B is often contracted by new users of heroin since needle-sharing frequently occurs. After contracting the disease, the second occurrence is rare. Thus, the number of hepatitis B cases may serve as an indicator of the number of heroin initiates (Schreeder 1978).

The number of hepatitis B cases per year has been reported by the Centers for Disease Control since 1966. A trend towards an increasing number of reported cases is thought to be due to several factors, among which are the improved reporting of hepatitis B cases, the increasing number of male homosexuals (a group thought to be especially susceptible to hepatitis B), and the linearly increasing trend of the number of intravenous cocaine users (Kozel et al. 1982).

Assuming that the improved reporting and the increase in the number of male homosexuals with hepatitis B and in the number of intravenous cocaine users follow a linear temporal trend, that trend can be removed. The residuals are thought to represent the number of hepatitis B cases associated with intravenous heroin use.

High School Senior Incidence Curve

High school senior heroin use is considered to be an incidence indicator because of the proximity of the year of first use to the year of the survey. Drug abuse treatment admission data from 1977-81 show that the mean age at first use for heroin is age 18 and that most users start between the ages of 16 and 20 (Johnston et al. 1982). Thus, a high school senior having used heroin is likely to have begun within a few years prior to the high school survey. Because of the close proximity of the year of first use to the date of the survey, the distribution of lifetime prevalence vs. the year of the survey can be considered an incidence curve with a 0 to 2-year lag. All high school data are taken from Johnston et al. 1982.

Heroin-Related Emergency Room Visits Reported to DAWN

The number of heroin-related emergency room visits is based on reports submitted to DAWN through 1982. Because DAWN is not a saturated sample of all hospitals in all cities in the United States, but is based on a percentage of the hospitals in 26 metropolitan areas plus a sample in other areas, a national projection is used (Hinkley and Greenwood 1982).

Heroin-Related Deaths Reported to DAWN

The number of heroin-related deaths reported to DAWN is also presented as a prevalence indicator. No attempt is made to project the number of heroin-related deaths to the nation, based on the number of cases occurring in the 26 DAWN metropolitan areas, because DAWN does not have a sample of medical examiners outside the major metropolitan areas.

Average Heroin Purity

Heroin purity as reported by the Drug Enforcement Administration (DEA) is found to be correlated with the number of emergencies and deaths related to heroin. Purity reported in the System to Retrieve Information from Drug Evidence (STRIDE) through 1978 and to the Domestic Monitor Program after 1978 was used.

FINDINGS AND DISCUSSION

The number of cases of heroin use combined from the three National Surveys on Drug Abuse is shown in table 1 by year of first use, 1965 through 1980. While the number of heroin users in any one survey is small, the pooled data produce frequencies large enough to establish trends. These data are not used to make estimates of the number of heroin initiates in any particular year, but are used to show a changing pattern over a several-year period.

TABLE 1. *Actual number of cases in pooled National Surveys on Drug Abuse (1977, 1979, 1982) reporting heroin use in lifetime by year of first use*

Year of First Use	Number of cases	Year of First Use	Number of Cases	Year of First Use	Number of Cases
1965	2	1970	24	1975	24
1966	2	1971	20	1976	28
1967	6	1972	23	1977	21
1968	13	1973	22	1978	20
1969	11	1974	26	1979	7
				1980	6

Figure 1 shows the population projections based on 2-year moving averages for all 3 surveys combined. The estimated number of initiates increased throughout the late 1960s and early 1970s from approximately 30,000 new users per year in 1965 to an average of approximately 180,000 new users per year between 1968 and 1972. Based on this analytic procedure, by 1978 and 1979, the estimated number of new cases declined to an average of approximately 80,000 per year.

The data in figure 1 show an elevated level in the estimated number of initiates between 1968 and 1976, with fluctuations between those years. Other data, such as heroin purity and treatment incidence curves, show the period between 1968 and 1976 actually consisted of two epidemics (DEA 1984; NIDA 1984). The two epidemic periods, i.e., the early 1970s and the mid-1970s, corresponded to the period of relatively plentiful heroin from the "French Connection" and Mexico. After approximately 1972, when the French Connection was broken, and after approximately 1976, when the Mexican poppy eradication program was fully operational, the heroin supply was reduced. The current supply appears to be from Southeast and Southwest Asia and Mexico, although the recent epidemic period is not evident in the household survey data.

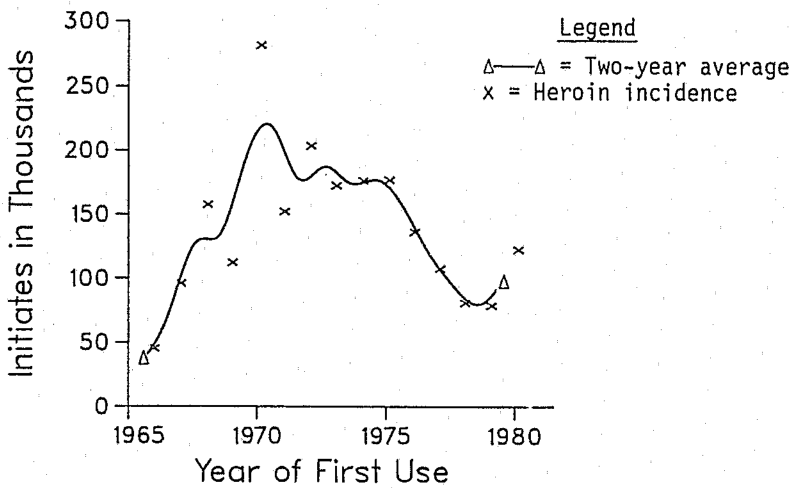


FIGURE 1. *Two-year moving averages of the estimated number of heroin initiates in thousands by year of first use*

Trends over time appear within subcategories of age and frequency of use. Figures 2A and 2B show the estimated number of new users by age groups 12 to 25 years, and 26 years and older. The number of initiates age 26 and older at the survey was highest in the early 1970s. The findings for persons age 26 and older at the survey (figure 2b) are consistent with the time series for treated initiates (figure 4), which shows a predominant peak in 1969 and smaller peaks thereafter (NIDA 1984). The average age of treated heroin users in 1981 was over age 26 (NIDA 1981).

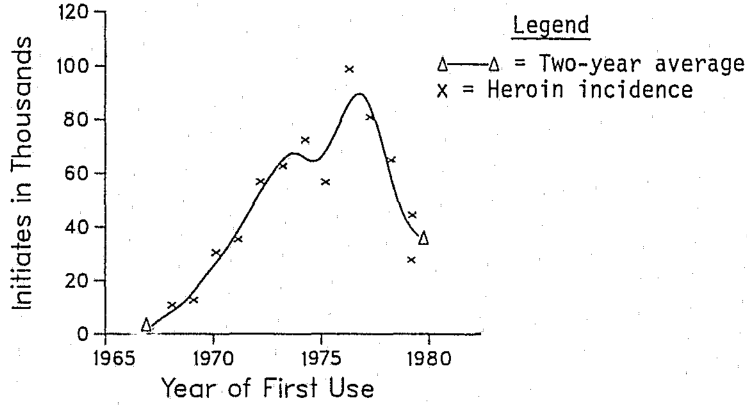


FIGURE 2A. *Two-year moving averages of the estimated number of heroin initiates in thousands by year of first use--ages 12-25 at the survey*

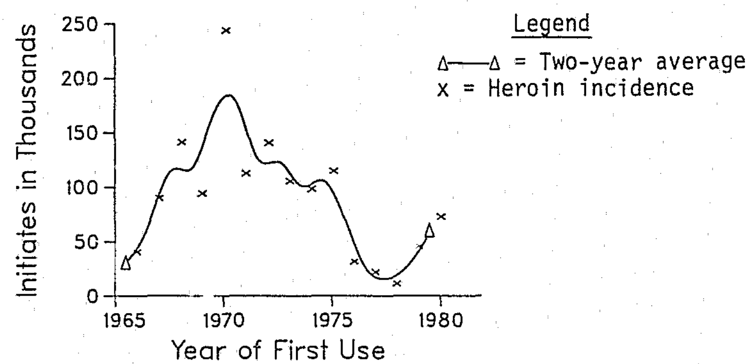


FIGURE 2B. *Two-year moving averages of the estimated number of heroin initiates in thousands by year of first use--age 26 and older at the survey*

The number of new users age 26 and older in households has shown a recent leveling or increasing trend. A similar recent leveling or increasing pattern can be seen for treated initiates by 1979 and 1980 (NIDA 1984).

The time series of young initiates shows the mid-1970s epidemic and no other. Persons ages 12 to 25 years interviewed in a survey conducted in 1977, 1979, or 1982 would be too young to be susceptible in the 1968 to 1972 epidemic period. Therefore, there appears to be an interaction between age and epidemic period.

Data from the three pooled National Household Surveys were used to examine the relationship between year of new, or first, use and "type" of heroin user. Two "types" of heroin users are defined. Persons who used heroin only one or two times in a lifetime were considered experimenters and those who used three or more times in a lifetime were considered continuing users. The estimated number of new heroin users in households by year of first use is shown in figure 3A for experimenters and in figure 3B for the continuing users.

Figure 3A shows the 1968-72 and the 1974-76 peaks of initiates. In 1974-76 the number of experimenters is greater than in 1968-72. In contrast, as shown in figure 3B, the number of initiates who became continuing users from 1968-72 is greater than from 1974-76.

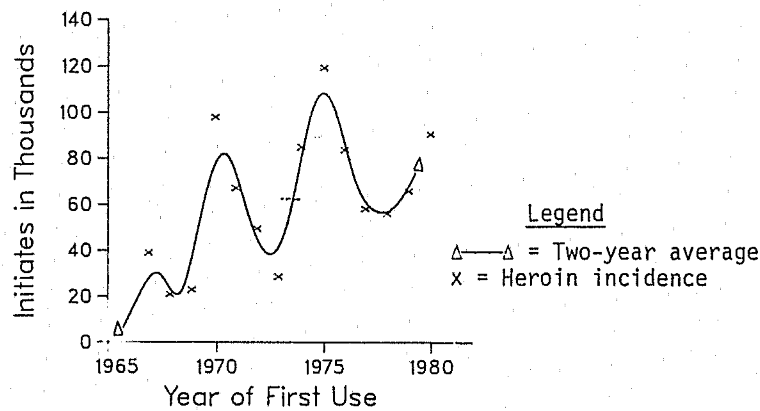


FIGURE 3A. *Two-year moving averages of the estimated number of heroin initiates in thousands by year of first use--for persons who used 1-2 times in lifetime*

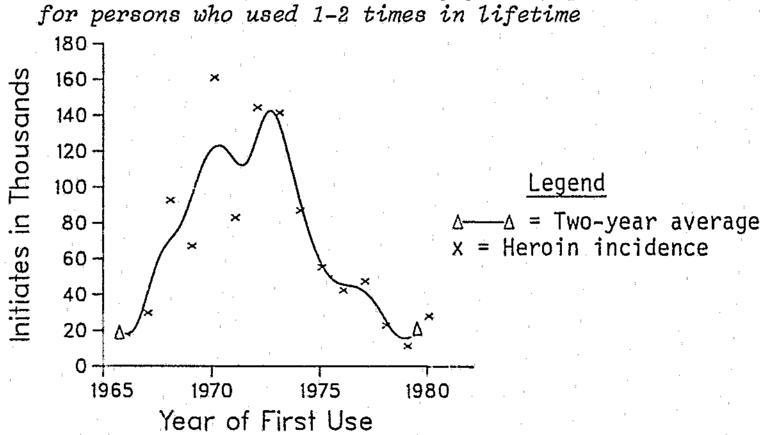


FIGURE 3B. *Two-year moving averages of the estimated number of heroin initiates in thousands by year of first use--for persons who used 3 or more times in lifetime*

Users who progress to frequent use of heroin usually do so within a short interval, 1 to 3 years, of their first exposure to heroin (Harding 1984). Therefore the distribution of frequency of use of heroin in 1974-76 as shown by the National Surveys is not a temporal artifact of the data collected.

The trends noted for epidemic periods of heroin initiation based on household survey self-reports were compared to trends based on periods of initiation reported by heroin users in treatment. The number of new heroin users who continue heroin use, in households, parallels the trends in the number of heroin users that eventually enter treatment. The estimated number of new heroin users entering a consistently reporting panel of federally funded treatment programs for the first time in a 5-year period by year of first use is shown in table 2 and figure 4. Table 2 and figure 4 show an epidemic period between 1968 and 1972. During the 1974-76 period the number of initiates was level compared to the earlier declining trend and was not as high as the 1968-72 peak. The larger 1968-72 peak is followed by a smaller 1974-76 epidemic period for initiates in treatment, a pattern similar to that noted for continuing heroin users.

The national distribution shown in figure 4 is the sum of two different regional distributions. The first is the late '60s and early '70s epidemic in most regions of the country and the second is the mid '70s epidemic occurring primarily in the western United States. The sum of the two makes the composite trend line appear level in figure 4 in the 1974-76 period.

TABLE 2. *Number of heroin users entering treatment by year of first use (corrected for treatment lag)*

Year of First Use	Number of Initiates	Year of First Use	Number of Initiates	Year of First Use	Number of Initiates
1965	4,920	1970	14,413	1975	10,357
1966	5,739	1971	12,584	1976	8.818
1967	8,459	1972	11,081	1977	7,310
1968	12,080	1973	10,864	1978	6,829
1969	15,085	1974	10,818	1979	7,519
				1980	8,080

SOURCE: National Institute on Drug Abuse. Client Oriented Data Acquisition Process. Data files for 1977 through 1981.

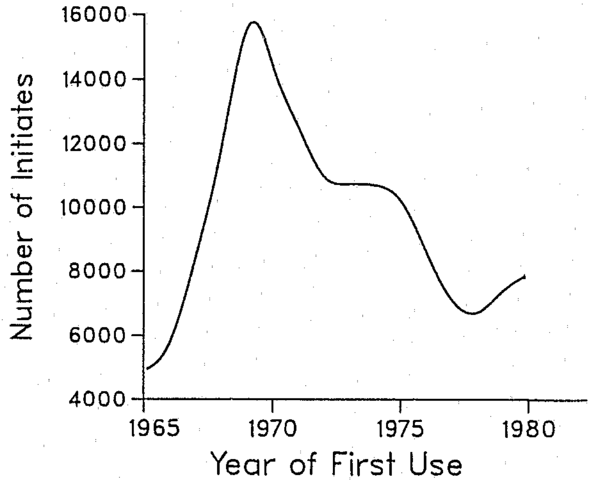


FIGURE 4. *Number of primary treatment admissions to federally funded programs by year of first use*

SOURCE: National Institute on Drug Abuse, Client Oriented Data Acquisition Process. 1977 through 1981 data files.

Using the function "number of household heroin initiates = B0 + B1 x the number of treatment heroin initiates" by year of first use, yields a B1 of 16.6 (F=38.13, n1=1, n2=13, p less than .001, R^2=0.74). Thus, it is estimated that 1 in 17 household resident heroin users will enter a panel of 402 consistently reporting federally funded treatment programs for the first time in a 5-year period (treatment data 1977 through 1981). The ratio of heroin initiates in households to all treated initiates would be much smaller.

The high R^2 value implies that the trends in periods of initiation based on self reports from continuing users in households were similar to those based on self-reports from users in treatment. Trends based on self-reported drug use were compared with other indicators, such as hepatitis B cases, high school survey data, heroin-related emergencies, heroin-related deaths, and average heroin purity.

The comparison of trends in heroin initiates to trends in hepatitis B cases required that a linear trend be removed. The linear trend is thought to be related to increasing numbers of male homosexuals, improved reporting, and an increase in the number of intravenous cocaine users. Table 3 and figure 5 show the residuals after the removal of the linear trend. The years in which the number of residual cases peaked were 1971 and 1977, approximately 1 to 2 years later than the year of the maximum number of initiates for household residents (i.e., in 1970 and 1975). Schreeder (1978) observed that 75 percent of the heroin users entering treatment contracted hepatitis B within 5 years of initiation. Thus, a lag between the number of household resident heroin initiates and a number of hepatitis B cases would be expected.

TABLE 3. *Residual number of hepatitis B cases after removal of linear trend*

Year	Residual Number of Cases	Year	Residual Number of Cases	Year	Residual Number of Cases
1965	-	1971	1580	1977	1579
1966	-416	1972	213	1978	-1448
1967	-667	1973	-1950	1979	-2224
1968	490	1974	-982	1980	125
1969	358	1975	294	1981	1050
1970	1546	1976	934	1982	217

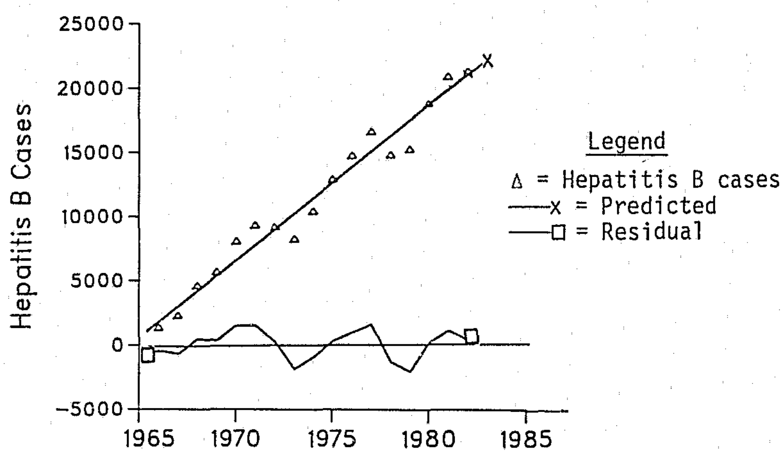


FIGURE 5. *Residual number of hepatitis B cases after removal of the linear trend*

SOURCE: Centers for Disease Control, 1975 through 1982.

The percent of high school seniors having ever used heroin, displayed in figure 6 by year of the survey, show a trend similar to that of other incidence indicators. In 1975, 2.2 percent of the high school seniors had ever used heroin. By 1980, the percent had declined to 1.1 percent. The relatively high prevalence among high school seniors in the mid-1970s and subsequent sharp decline correspond to the decline in incidence shown in the national household survey data for 12- to 25-year-olds.

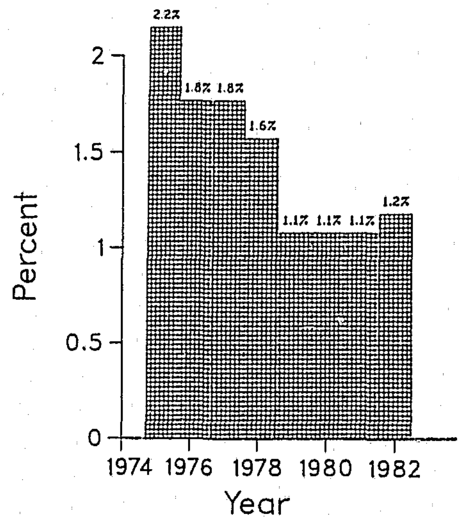


FIGURE 6. *Percent of high school seniors using heroin at least once in lifetime*

SOURCE: Johnston et al. 1983, p. 32.

Figure 7 shows the number of heroin-related emergency room visits by month from 1973 through 1982 reported to DAWN and projected to the nation. These data are based on a national projection from DAWN data for that period (Hinkley and Greenwood 1982). In figure 7, the mid-1970s epidemic period can be seen, along with the stabilizing trend in the late 1970s after a decline following 1976. Recent increases can be seen in the emergency room data, although these data are too recent to allow comparison to national survey data.

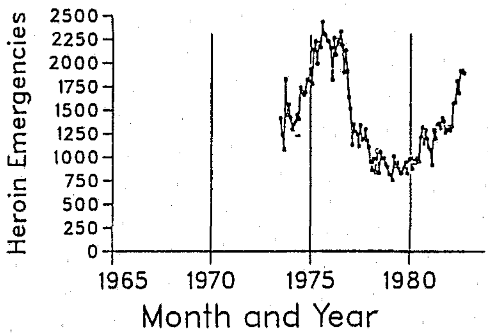


FIGURE 7. *Number of heroin-related emergency room visits projected to Nation*

SOURCE: National Institute on Drug Abuse 1984, p. 9.

Figure 8 shows the number of heroin-related deaths reported to DAWN. The data show the mid-1970s epidemic period and the recent increase. The peak in the 1974 through 1976 period corresponds to the increase in the number of heroin initiates in that period.

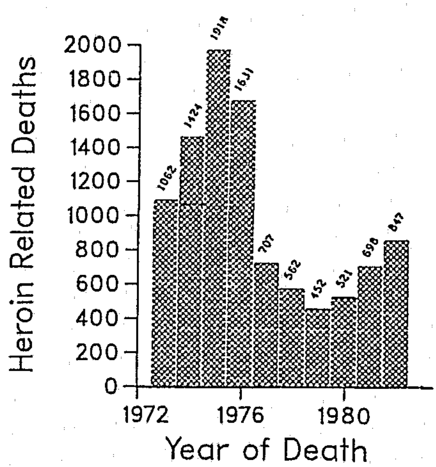


FIGURE 8. *Number of heroin-related deaths reported to DAWN*

SOURCE: National Institute on Drug Abuse 1984b.

Average heroin purity January 1973 through September 1983, based on data provided by the DEA (1983), is shown in figure 9. As was noted for the other heroin indicators, the purity was elevated during the mid-1970s epidemic period, declined between 1976 and 1979, and shows a recent increase.

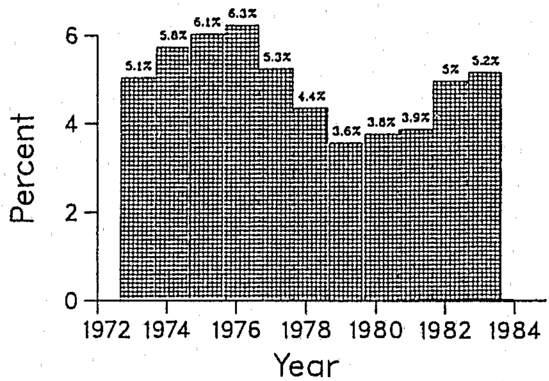


FIGURE 9. *Average heroin purity reported to STRIDE and Domestic Monitor Program*

SOURCE: Drug Enforcement Administration 1984.

SUMMARY

Because of the small proportion of the population reporting ever having used heroin, the year of first use data from NIDA's National Household Surveys on Drug Abuse conducted in 1977, 1979, and 1982 were pooled to show the number of new users in the household population by year of first use. In addition, the data were "smoothed" by using a 2-year moving average. The early 1970s and the mid-1970s epidemics were evident. These epidemic periods occurred at the same periods reported by high school seniors and by heroin users in treatment. The household self-report data trends based on age and frequency of use were also consistent with the trends in periods of initiation reported by heroin users in treatment as noted in drug abuse treatment admission data for year of first heroin use.

Trends in indicators of heroin epidemics were compared with trends based on self-report data from the National Household Surveys. The trends in hepatitis B cases, heroin-related emergency room visits, heroin-related deaths, and the average retail heroin purity were consistent with the epidemic periods suggested by the household data.

This consistency among the three sources of self-reported data on trends in year of first heroin use combined with the consistency of these self-reported data with the trends based on the indicators of heroin epidemics offers some validation to the use of retrospective direct questions concerning age of first use of heroin to monitor heroin incidence in the household population.

REFERENCES

Centers for Disease Control. Mortality and Morbidity Weekly Report Annual Summary 1975, 1976, 1977, 1978, 1979, 1980, 1981. Data for 1982 and 1983 were from computer printouts that form the basis of the Mortality and Morbidity Weekly Report Annual Summary.

Drug Enforcement Administration. Domestic Monitor Program, July-September 1983. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1984. 34 pp.

Fishburn, P.M. Heroin estimator development: The nominative technique for heroin prevalence estimation. In: Rittenhouse, J.D., ed. Development Papers: Attempts to Improve the Measurement of Heroin Use in the National Survey. National Institute on Drug Abuse Administrative Report, 1980. 170 pp.

Harding, W.M. Can compulsive opiate users establish stable patterns of controlled use? In: Epidemiology of Drug Abuse: Research and Treatment Issues. Volume II. Community Epidemiology Work Group Proceedings. Rockville, Md.: National Institute on Drug Abuse, June 1984. pp. 172-190.

Hinkley, S., and Greenwood, J. Computer Printouts, National Projections of Emergency Room Visits, data through September 1982. Washington, D.C., Drug Enforcement Administration, 1983.

Integrated Software Systems Corporation. TELLAGRAF Version 4.5. San Diego, 1983. 40 pp.

Johnston, L.D.; Bachman, J.G.; and O'Malley, P.M. Student Drug Use, Attitudes, and Beliefs, National Trends 1975-1982. Prepared for the National Institute on Drug Abuse by the University of Michigan Institute for Social Research. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1982. 134 pp.

Kozel, N.J.; Crider R. A.; and Adams, E.H. National surveillance of cocaine use and related health consequences. Morbidity and Mortality Weekly Report, 31(20):265-273, 1982.

National Institute on Drug Abuse. Client Oriented Data Acquisition Process. Data files for 1977 through 1981.

National Institute on Drug Abuse. Drug Abuse Warning Network. Data files through 1982.

National Institute on Drug Abuse. Heroin Work Group Preliminary Report, Rockville, Md., 1984a.

National Institute on Drug Abuse. Heroin Work Group Report, Rockville, Md., 1984b.

National Institute on Drug Abuse. National Survey on Drug Abuse. Data files for 1977, 1979, and 1982.

National Institute on Drug Abuse. Statistical Series, Trend Report January 1978-September 1981. data from the Client Oriented data Acquisition Process (CODAP). Rockville, Md., Series E, Number 24, 1981. p. 36.
Schreeder, M.T. CDC hepatitis laboratory presentation. In: Proceedings Community Correspondents Group Meeting Five. National Institute on Drug Abuse Research Monograph. Washington, D.C., 1978. pp. 35-38.
Sirken, M.G. Network surveys of race and sensitive conditions. In: Advances in Health Survey Research Methods. NCHSR Research Proceedings Series, Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1975. 15 pp.

AUTHOR

Raquel A. Crider, Ph.D.
Division of Epidemiology and
Statistical Analysis
National Institute on Drug Abuse
5600 Fishers Lane
Rockville, MD 20857

ESTIMATING HEROIN IMPORTS INTO THE UNITED STATES

Keith L. Gardiner
Raymond C. Shreckengost

INTRODUCTION

One problem that plagues strategic narcotics analysis--analysis of major trends in the illegal production, trafficking, and consumption of assorted dangerous drugs--is the poor quality of data or lack of it on crucial aspects of what is grown, manufactured, moved, and consumed in the international illegal drug world. There is no fully adequate substitute for data, but there are useful analytical aids that help get around the problem by making better use of the data that are available.

This paper describes the development of one such aid, a dynamic simulation model implemented on a computer. The method used in designing the model, System Dynamics, has several features that are highly advantageous in this sort of analysis--including, for example, the use of expert opinion to identify the critical factors that influence the behavior of the system and to see how a change in one factor affects the others (Forrester 1961, 1975; Roberts 1978).

This paper illustrates how the fundamental factors affecting the behavior of the heroin system are interrelated, and it provides several examples of how estimates, or predictions, produced by the model have spotlighted erroneous data, supplied missing data, and anticipated data that become available only later. Most important, the model provides a general structure not necessarily restricted to heroin; it may be applicable to other illegal drug systems.

Model development was initiated during the summer of 1983 to reduce the time required to estimate the amounts of heroin being imported into the United States from foreign sources. Data supplied by the Office of Intelligence of the Drug Enforcement Administration (DEA) and the National Institute on Drug Abuse (NIDA) were essential to the development of this model (Greenwood and Crider 1978; Rosenquist 1983).

The views expressed in this paper are those of the authors and do not necessarily reflect those of the Central Intelligence Agency.

THE MODEL AND HOW IT WORKS

The key to the design of the model was the notion that the behavior of the heroin system in the United States would be affected most importantly by how much heroin is available at any time (Inventory) in comparison to the amount the addicts want (Desired Inventory). This ratio is referred to in the model as the Relative Abundance Measure. This indicates the surplus, adequacy, or shortage of the heroin supply at any time. This, in turn, affects directly such things as the price of heroin, its purity, and the number of heroin users. These influences are depicted in figure 1.

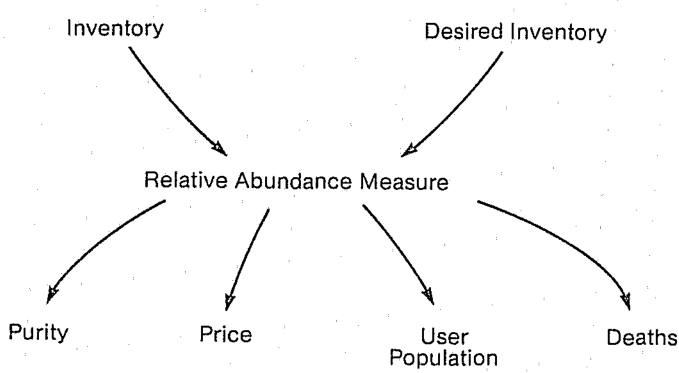


FIGURE 1. *Supply ratio influences in the heroin system, United States*

As is often the case, these relationships appear simple and obvious once described, but they are not so apparent beforehand. In addition, to get to the essence of what makes the system work, some of the clutter of conflicting data must be resolved. In this model only one factor was troublesome--the definition of a heroin user. There are heavily addicted users (who spend substantial amounts of time in jail); light, sporadic users; and moderate users who fall between these extremes. The heavy users are comparatively few, the occasional users comparatively many. For the purposes of this model, using NIDA data, we created an "average" user who consumes a package of heroin a day that contains 22.6 milligrams of pure heroin when the United States inventory is at the desired level. As will be seen later, the purity of the heroin consumed and the number of users both increase when supply increases more than consumption, and they decrease when there is a shortage of heroin. Total consumption is thus affected by both the number of users at any time and the purity of the heroin in the package they buy.

In the heroin problem, as in most other social problems, the important factors in the system are numerous and highly interrelated, and this makes analysis based on pure reasoning (that is, without the use of analytical aids) difficult, tiring, uncertain, and almost impossible to retrace. One function of the model is to make the relationships clear and explicit, and then let a computer perform the tedious processing. We do this by creating a "decision rule" about, for example, how much on the average the purity in a package of heroin will rise as the available quantity of heroin increases; we express this rule as a mathematical reference in the model, and then let the computer calculate purities as the imports of heroin rise and fall. The multiple effects of the factors (variables) influencing the behavior of this model are shown in figure 2.

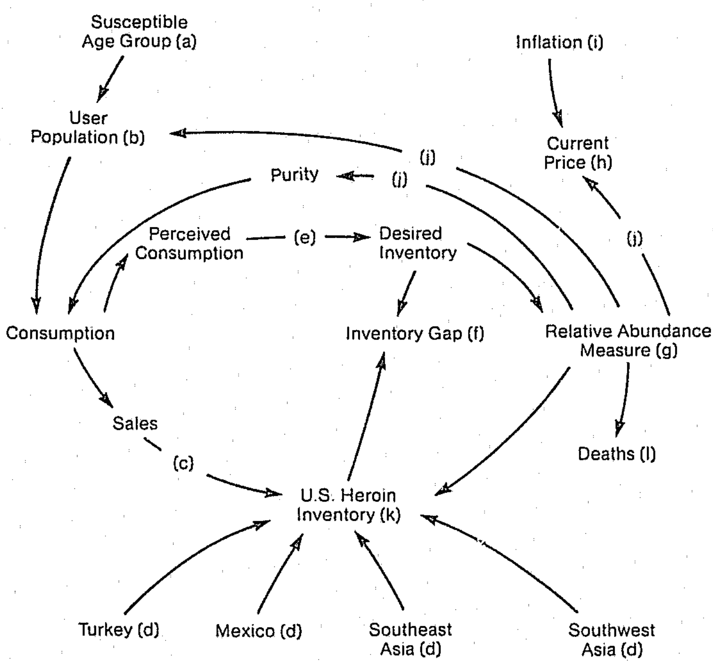


FIGURE 2. *Complete influence diagram of the heroin system, United States*

Notes on the Model

Details of our assumptions about the variables in the heroin system model are as follows (letters below correspond to those in figure 2):

a. At the upper left-hand corner is the Susceptible Age Group. This group consists of the 14- to 34-year-olds in the population, as derived from census data. The User Population is a fraction of this group, which varies with the availability of heroin--the Relative Abundance Measure. If desired, the User Population could be shown as responding to Price, in which case Price would reflect supply/demand and the User Population would change with Price. As will be explained later, Price is not directly affected as it is in normal marketing systems, and the Relative Abundance Measure was selected as the more appropriate factor to use.

b. The size of the User Population, and the Purity of what addicts buy, determine Consumption. The various classes of users and the amounts they normally use per day are not treated separately, but they could be if desired.

c. Sales to support Consumption reduce the U.S. Heroin Inventory.

d. The four sources of heroin imports since 1973 for the United States heroin market are arrayed at the bottom of the figure in the order in which they first delivered supplies to the United States.

e. The Desired Inventory is taken as five times the weekly Perceived Consumption. This provides a slight buffer against small variations in overall supply, but holds down risks of loss. Perceived Consumption serves to represent the lag in the response of the operators in the system to the ups and downs in Consumption in the short term.

f. The Inventory Gap is merely the difference between the Desired Inventory and the actual inventory, the U.S. Heroin Inventory.

g. The Relative Abundance Measure provides a single index for the actual/desired inventory status. This ratio affects both the heroin User Population and the Purity of the heroin it buys.

h. As distinct from purchases of other things, such as food, the addict does not buy heroin directly at so many dollars per milligram but buys instead a package in which the amount of pure heroin is small and variable. Thus, price per milligram is indirectly derived.

i. In the model, Price (per milligram pure) is related to the Relative Abundance Measure and Inflation. The values for inflation are taken from the Commodity Price Inflation Index.

j. The relationships among the Relative Abundance Measure and Purity, Price, and the User Population are of particular interest. Purity and the User Population are parts of loops--that is, the Relative Abundance Measure affects Consumption, which affects the U.S. Heroin Inventory, which affects the Relative Abundance Measure, which affects Purity, and so on. A similar situation exists with the User Population: changes in Purity and User Population affect Consumption and the U.S. Heroin Inventory, and these are influenced by the Relative Abundance Measure.

k. Over the 10-year span, the ratio of the U.S. Heroin Inventory to the Desired Inventory ranges from a low of .8 (undersupply) to 1.8 (oversupply). Purity bottoms out at about 3.5 percent when it is in short supply, rises rapidly until the inventory overage amounts to about 50 percent, and then begins to taper off slightly. The variation in User Population is much less dramatic, falling a bit when heroin is in short supply and increasing significantly only when the excess is greater than 20 percent. Prices come down as heroin becomes more abundant, but seem to bottom out at $0.90 per milligram of pure heroin.

l. As described in this paper, the effect of the Relative Abundance Measure on Deaths is more complex than the effect of the ratio on Purity, Price, and User Population. The system exhibits two modes of behavior, one when the ratio is rising and the other when it is falling.

Considering these influences, and given the heroin imports described below, the model forecasts what the User Population, Purity, Prices, and Deaths will be. These predictions are generally referred to as the model's behavior, and the way the model behaves is determined by the influences, or structure, described above. Whether the model can be used confidently is largely dependent on how closely the predictions of the model match their real-life counterparts: for example, at any given moment, do the model's predictions for Purity match independently measured national averages for purity? The independently measured values for purity, price, and heroin-related deaths are not direct inputs for the model. They are used only for comparison with the model-generated estimates to assist in evaluating how well the model behavior matches real-life behavior. The particular question here is whether the model will provide good predictions using only the fluctuating heroin imports over the last 10 years as the input to the model.

MODEL INPUTS

The sources and amounts of heroin imports used in the model correspond, with a single exception discussed later, to the sources and amounts commonly accepted by the various agencies concerned with drug abuse. These inputs are shown in tabular form in table 1 and graphically for the same 10-year period in figure 3.

TABLE 1. *United States imports of heroin in pure metric tons*

Country of Origin	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983*
Turkey	1.0	0.5									
Mexico	3.7	5.0	5.6	4.0	3.1	1.9	1.0	1.7	1.8	1.8	—
Southeast Asia			1.0	2.0	2.0	1.7	1.2	0.4	0.4	0.7	—
Southwest Asia					0.4	0.8	1.4	2.1	2.2	2.7	—
Total	4.7	5.5	6.6	6.0	5.5	4.4	3.6	4.2	4.4	5.2	5.4

*The 1983 estimate procedure is described later.

The total heroin imports are plotted again in figure 4, which shows how these imports relate to the Relative Abundance Measure. Using expert opinion, we built into our model the assumption that heroin traffickers attempt to hold 5 weeks' supply (at the current consumption rate) as a buffer against surges in demand and delays in supply. This provides a degree of inertia in the system and causes changes in the Relative Abundance Measure to lag behind the changes in Imports. The relatively high amount of heroin on hand when the model begins in 1973 results from a sustained growth in heroin imports from Mexico prior to 1973.

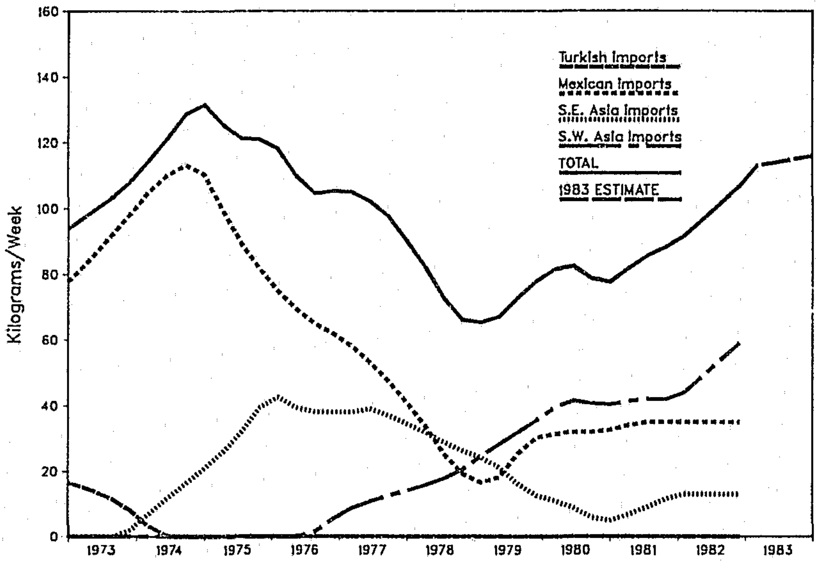


FIGURE 3. *Imports*

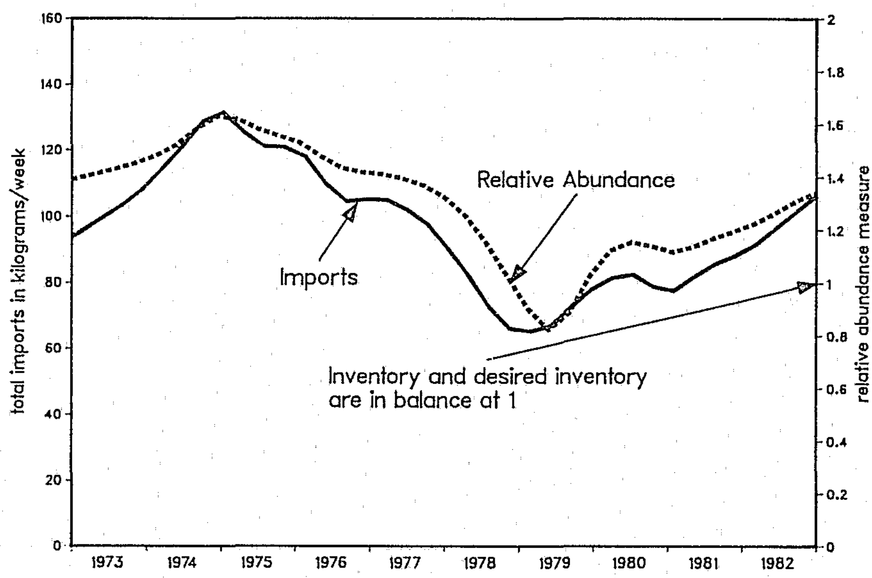


FIGURE 4. *Imports and relative abundance measures*

MODEL VALIDITY

There are no absolutely valid models because such models would have to replicate any "real-world" system faithfully and exhaustively in even the most trivial detail; all models, including all mental models, are thus simplifications. However, as described earlier (Shreckengost, this volume), if the key factors influencing the system are identified and correctly interrelated, the behavior, or predictions, of the model will parallel the behavior of the real system closely enough to validate, or provide confidence in, the model. In addition, confidence in the accuracy of the model is enhanced when its behavior, using inputs from the real world, does not violate common sense. The following discussion deals with one dimension of the model--its ability to predict the size of the addict population--for which there are no usable matching independent data by which to show the model's validity, but where the value of the model is demonstrated by its ability to spot a possible error in the data supplied to the model. Three areas in which data are available for comparison with numbers the model generates--Purity, Price, and Deaths--are then discussed. By matching the figures that the model generates on these three variables with the independently measured numbers, we can determine how accurately the behavior of the model parallels the behavior of the real system.

User Population

Figure 5 shows the decision rule stating how the Relative Abundance Measure affects the size of the User Population. The effect of the Relative Abundance Measure on the User Population is quite constant when the Inventory and Desired Inventory are nearly balanced, but lowers the population slightly when shortages occur and increases the population slowly with increases in supply--that is, an 80 percent oversupply produces only a 10 percent rise in the population. Our model shows only how changes in the Relative Abundance Measure affect the percentage of change in the User Population. The actual number of addicts added or subtracted depends on the size of the 14- to 34-year-old age group. As the size of this group increases, logically the user population will also increase for any given Relative Abundance Measure.

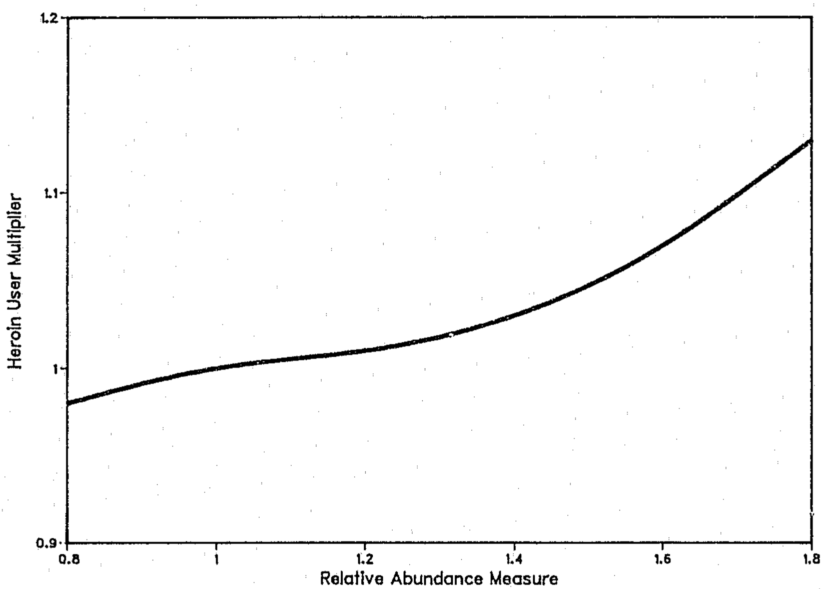


FIGURE 5. *Effect of relative abundance on user population*

The addict population numbers the model produces do not violate common sense and are consistent with other estimates available from various surveys and analyses. However, there are no good data on the actual size of the addict population in the United States. Consequently, the validity of the model cannot be checked by comparing the numbers it produces with independent user population measurements. The model-estimated User Population is shown in figure 6. As a baseline from which to observe fluctuations in the size of the addict population that are independent of the normal growth that occurs as the 14- to 34-year-old age group increases, we derived an estimate using population data that the addict population equals 0.65 percent of the 14- to 34-year-old population when supply and consumption are in balance (Greenwood and Crider 1978). The natural demographic increase in that percentage is shown in the figure as a "reference group" against which the size of the addict population predicted by the model rises and falls in relationship to how much heroin is available. The size of the population swelled with the growing availability of heroin through 1975, and then remained fairly constant through 1978 even though the amount of heroin available dropped, because the growing age group population offset the decline in the percentage of the age group addicted to heroin.

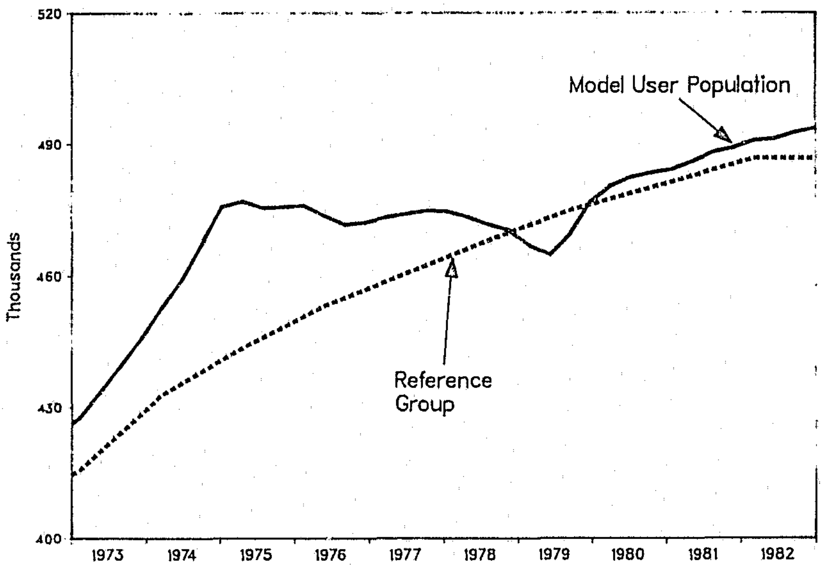


FIGURE 6. *Comparison of growth in heroin user population with growth in reference population (14-34-year-old age group)*

Even though we cannot independently verify that the model's user population projections are accurate, we were able to use them to reveal a logical inconsistency between one of the import figures used as an input to the model and the size of the user population that the model indicates would be required to consume that amount of heroin. Originally, the estimate for heroin imports for 1975 supplied to the model included 6.5 tons from Mexico instead of the 5.6 tons shown in table 1. This was inconsistent with the size of the User Population predicted by the model for that year. In particular, the rate of increase in the size of the population would have had to have jumped considerably to produce a User Population large enough to consume so much heroin. This would run contrary to the commonsense notion that, since there is probably an upper limit on the fraction of the total population that might become heroin users (as there is for cigarette smokers), the rate of increase should tend to subside as heroin becomes more abundant. In addition, that much surplus heroin would have had effects on Purity and Price that were not borne out by the data available for that year. Because of these inconsistencies, we double-checked the import figures and discovered that 5.6 tons provides a better "fit" and is consistent with other independently available information. This is one illustration of the way in which dynamic simulation models can highlight logical inconsistencies, which may not be apparent until otherwise plausible data are viewed against a comprehensive and consistent framework.

Heroin Purity

The way the Relative Abundance Measure affects Purity in the model is shown in figure 7. When heroin is in relatively short supply, the amount of pure heroin in the package the user buys is also low. When heroin supplies are abundant, the amount of pure heroin increases. Unlike the User Population, Purity changes quite dramatically with changes in the Relative Abundance Measure. As with the User Population, however, the amounts of change are not constant but depend on the adequacy of heroin supply that exists at any particular time.

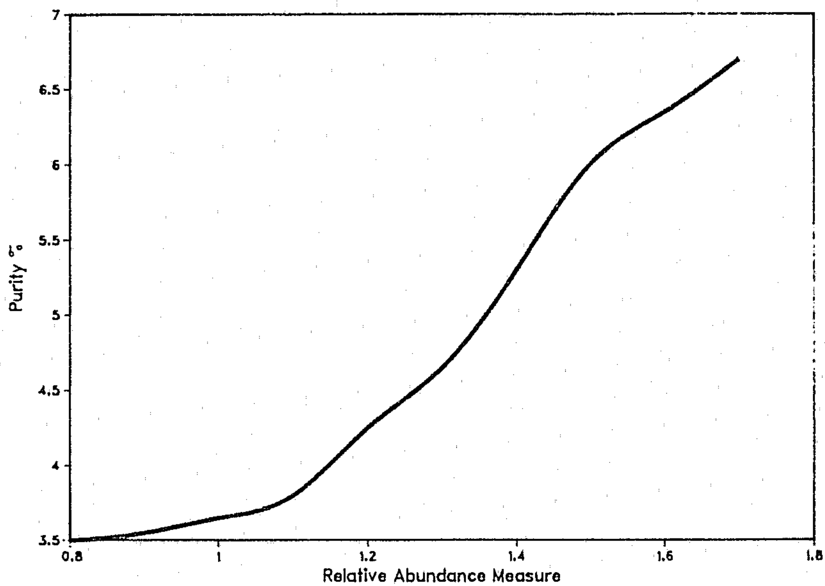


FIGURE 7. *Effect of relative abundance on purity*

Samples of heroin have been analyzed by the DEA for purity over the years, and the comparison of these data with the model-generated Purity values is shown in figure 8. In view of the simplicity of the model and the sampling and statistical problems involved in determining purity, the correspondence between the purity measured in the heroin samples and the Purity estimated by the model over the 10-year period is surprisingly good. Only annual average purity values are used for 1973, 1974, and 1975, after which quarterly values are available. Since Purity figures in the model are derived in part from known imports, and because the correspondence between sample and model Purity data is so close, we can reverse the procedure and use known purity figures to estimate what imports were at a given time. We have used this procedure three times to produce "missing" import figures. In the first instance, we used sample purity data from earlier years to estimate what the residual imports from Turkey must have been in the 1973-74 period following the halt in Turkish production in 1972. We have no figures from other sources on Turkish imports for these years. With data only for Mexican imports, the model-generated Purity values fall well below the 5.1 percent and 5.8 percent found in sample data. The amount of imports from Turkey that we have inserted in table 1 and figure 3 are those required to cause model Purity to be consistent with the observed system values. These imports are intuitively reasonable and "fill in the blanks."

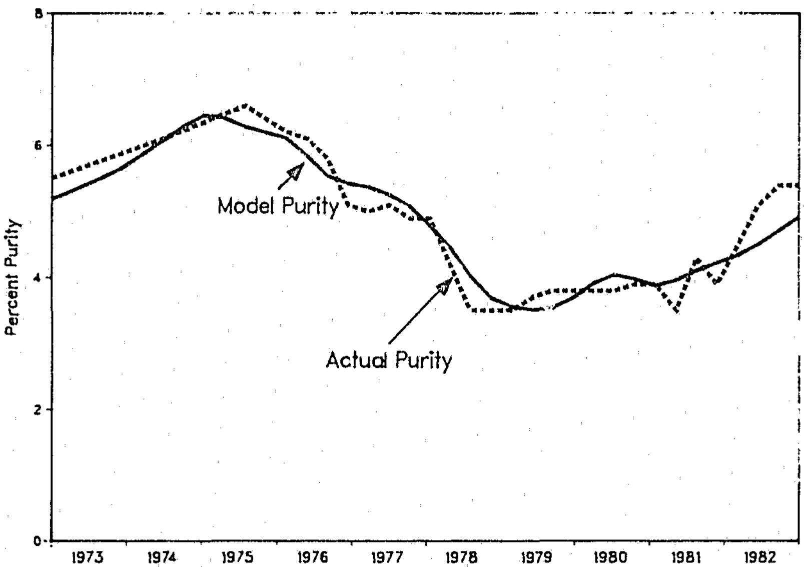


FIGURE 8. *Actual purity vs. model purity*

The second case in which we used purity figures to derive imports was in attempting to estimate overall imports for 1982. Our original information was that imports for 1982 were about the same as in 1981. This, however, resulted in model Purity values for 1982 that were quite a bit lower than the measured purity values. Here, again, comparison of model and sampled purity suggested that the imports were understated. Subsequently, estimates of the imports from Southeast and Southwest Asia were revised upwards by the DEA. While this brings the model and real-world data into good correspondence, the total imports still appear to be slightly underestimated.

In addition to using purity figures to double-check available import statistics, we have used them to produce an estimate of how much heroin was imported into the United States in 1983. Using the DEA-measured values for purity for 1983, the model indicates that some 5.4 metric tons of heroin would have had to be imported into the United States in that year to generate model Purity values that correspond to the measured numbers. To check the validity of this import figure, we also generated model Prices for 1983 to see if they would correspond to available measured street prices. The correspondence was good: street prices for the first three quarters of 1983 (the fourth-quarter prices are not yet available) vary from $2.28 to $2.43 per milligram, and the model Prices range from $2.22 to $2.45.

Heroin Prices

Price moves opposite to Purity with changes in the Relative Abundance Measure. Also, as shown in figure 9, the prices fall in an almost straight line as the Relative Abundance Measure rises, but then level off rather suddenly (in constant 1970 dollars) as a price floor is reached below which it does not make economic sense for a trafficker to continue to deal in heroin (in part because of the personal risk involved).

This lower limit on prices influences prices strongly during the buildup of imports, and accompanying surpluses, into 1975. With a growing surplus of heroin supply over consumption, a continuing drop in price might have been expected, but instead the price floor was reached and, in fact, nominal prices rose slightly because of inflation. Notice, as shown in figure 10, that the model Price estimates are higher than the observed prices in 1982--another indicator that the estimated imports may be low.

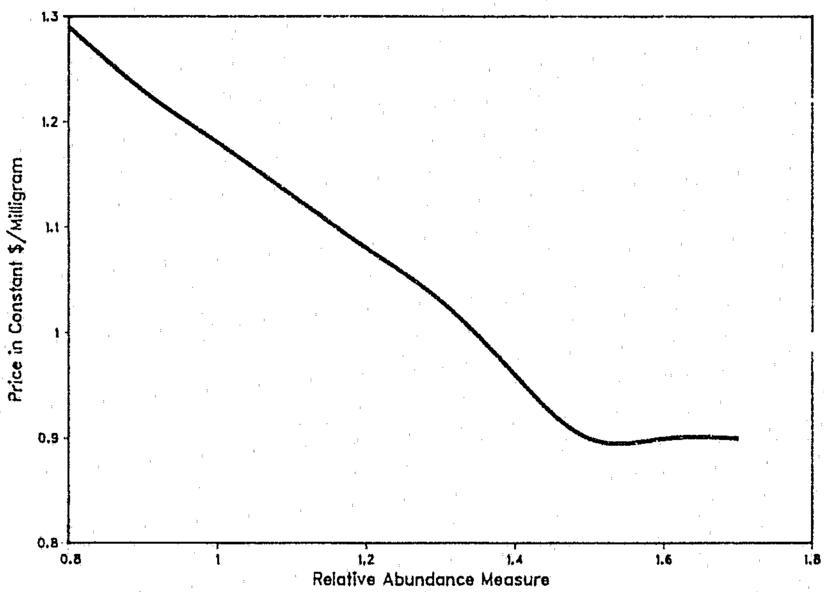


FIGURE 9. *Effect of relative abundance on price*

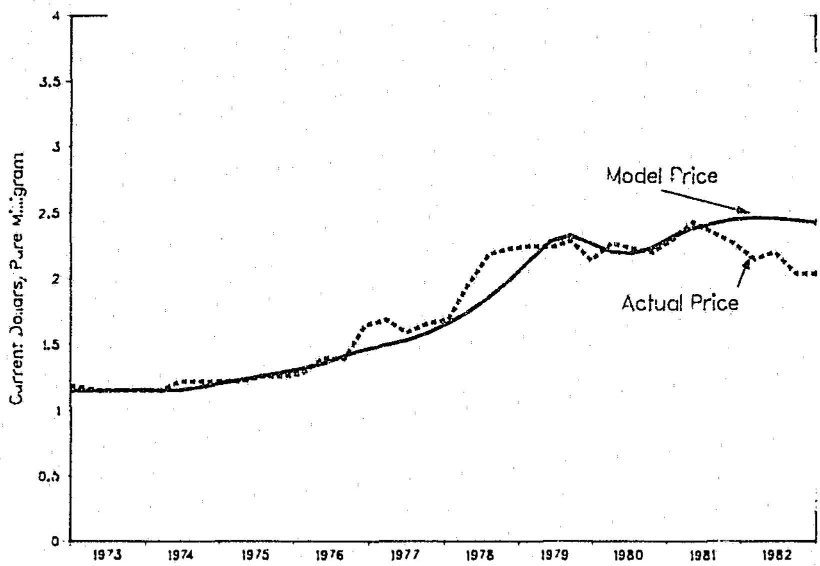


FIGURE 10. *Comparison of model price and actual price*

Heroin-Related Deaths

Figure 11 shows how the relative abundance or shortage of heroin affects the number of heroin-related deaths. It is more complex than the Purity and Price relationships because two distinct modes operate--one when the amount of heroin is increasing and the other when the supply is falling. This is required because when the supply of heroin is falling, the death rate does not retreat along the same path that it followed when supplies were increasing. Use of only the rising relationship results in major departures between model-produced Death estimates and the measured values. These two modes are believed to reflect the effect of Purity changes as the Relative Abundance Measure rises and falls. When supply is increasing, purity is also increasing, and the addict is subject to unexpectedly high doses, which may be fatal. On the other hand, when supplies and purity are falling, the likelihood of overdosing, either accidentally or deliberately, is reduced.

Figure 12 compares the heroin deaths reported from 1973 through 1981 with the model's predictions. It is important to appreciate that the definition of what constitutes heroin-caused or heroin-related deaths is neither precise nor uniform in the United States. Contributing factors, such as alcohol, are often present, so the need for subjective judgments is understandable and unavoidable.

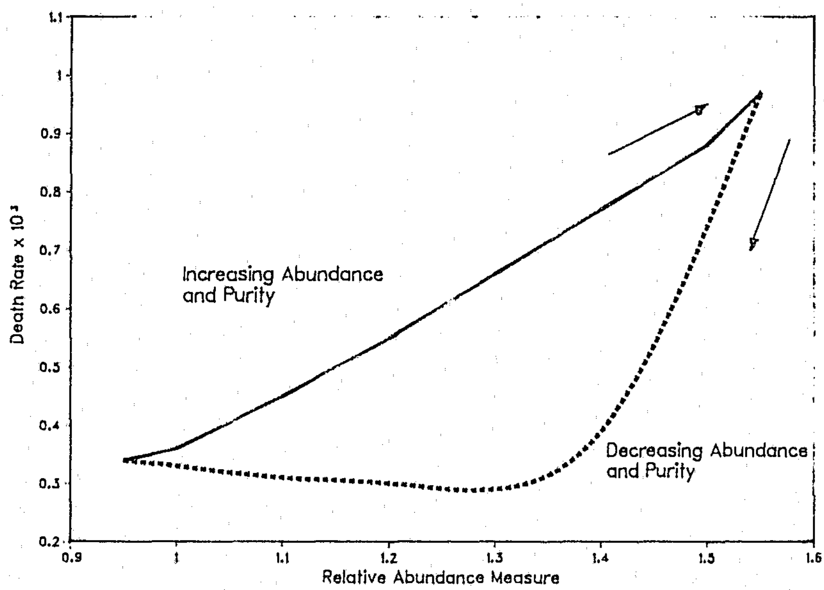


FIGURE 11. *Effect of relative abundance on deaths*

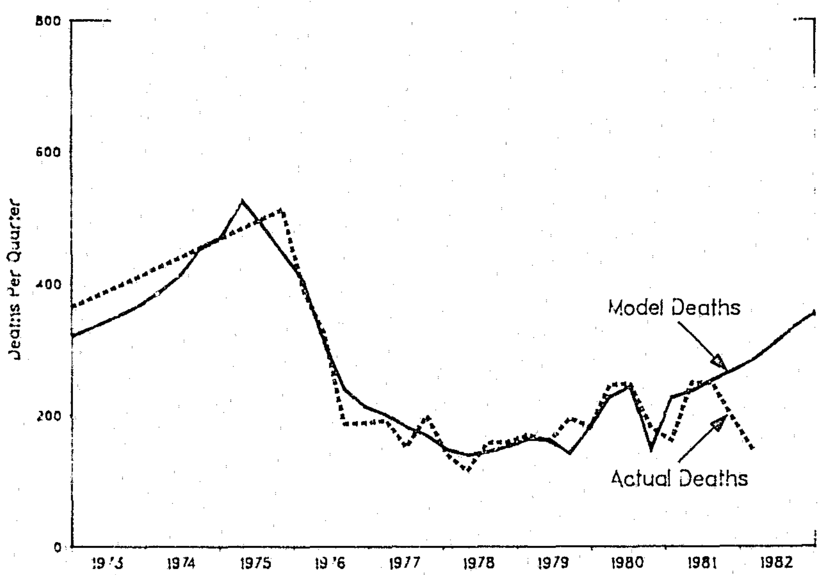


FIGURE 12. *Comparison of model deaths and actual deaths*

The reason 1982 sample data are incomplete illustrates, in part, the problems encountered in obtaining "hard data." For example, the number of heroin-related deaths in New York City is problematic because such deaths are recognized only when confirmed by a toxicological examination. Even if a victim is found in surroundings where the cause of death seems apparent--for example, heroin scattered about and a syringe in hand--if a toxicological examination is not done, the death is not tallied as heroin-related. The model indicates, however, that Deaths probably rose in 1982; we await confirming data.

IMPLICATIONS AND POSSIBLE FUTURE APPLICATIONS

This modeling effort is still in a developmental phase. We need additional time and research to determine how robust it is and whether it can be applied with equal facility to other aspects of the heroin market in the United States, whether it can be extended to non-United States markets, and whether it can be adapted to deal with illegal drugs other than heroin. In pursuing possible future applications of this model to the United States heroin market, of the various relationships demonstrated among elements of the heroin situation by this model, the one that stands out as potentially most exploitable involves purity. The notion that purity measurements of seized or purchased samples tend to rise and fall as heroin supplies grow and shrink has long been known. However, the relative precision with which the model predicted changes in heroin purity over a 10-year period indicates a relationship between Purity and such other factors as Imports and Consumption that is sufficiently strong to suggest the use of purity figures alone as a powerful, timely indicator of the state of heroin abundance in the United States at any given moment. In particular, it may be possible to use purity figures to estimate whether, and how much, imports of heroin and the size of the United States heroin addict population are expanding, holding steady, or declining. The advantage of using Purity as an indicator for these other factors is that purity measurements are usually available on a fairly "real-time" basis.

One problem with using purity measurements other than for what they intuitively suggest is that purity often varies sharply among daily samples obtained in different regions of the United States and among individual samples within a region. The large variations in purity at the local level have tended to mask the relationship with the other dimensions of heroin supply and consumption that are revealed when averages are used. The model uses national averages, thus smoothing out the momentary peaks and valleys and providing data that can be compared with other aggregate figures such as total imports and changes in the national user population.

Another possible use for the model would be to apply the same basic structure that was used on national level factors to monitor heroin supply and consumption at the regional/city level of the heroin system. If, for example, variations in purity from region to region were detected and the source countries were known, it might be possible to define and monitor lines of supply and communication in the national system. The time element involved in moving heroin supplies within the United States could be explicitly represented in a national heroin model, thereby adding to the possibility of predicting changes in the system.

A further interesting application of the model would be to attempt to use it to understand supply and demand relationships for other illegal drugs. To the extent that such drugs as cocaine, marijuana, and certain synthetics follow the same basic dynamics as heroin supply and demand, they could be modeled in the same fashion. Much more ambitious, but theoretically equally feasible, would be to design and develop a comprehensive, integrated national drug model, using the System Dynamics "loom" to weave the various individual drug systems together. A key advantage that might be afforded by such a model would be a better understanding of the addicts who use multiple drugs--the polyusers--and the tradeoffs that might occur in concentrating intelligence and law enforcement resources against particular drugs in the system.

REFERENCES

Forrester, J.W. Industrial Dynamics. Cambridge, Mass.: MIT Press, 1961. 464 pp.

Forrester, J.W. Collected Papers of Jay W. Forrester. Cambridge, Mass.: MIT Press, 1975. 284 pp.

Greenwood, J., and Crider, R. Estimated Number of Heroin Addicts, 1977. National Institute on Drug Abuse. Rockville, Md.: 1978.

Roberts, E.B., ed. Managerial Applications of System Dynamics. Cambridge, Mass.: MIT Press, 1978. 699 pp.

Rosenquist, E. Private communication. 1983.

AUTHORS

Keith L. Gardiner, M.A.
Raymond C. Shreckengost, M.S., M.A., M.P.A.

Central Intelligence Agency
Washington, DC 20505

ESTIMATING THE SIZE OF A HEROIN-ABUSING POPULATION USING MULTIPLE-RECAPTURE CENSUS

J. Arthur Woodward, Douglas G. Bonett,
M.L. Brecht

INTRODUCTION

During the last 10 years, the Federal Government has attempted to monitor trends in heroin abuse prevalence to develop appropriate drug prevention and control activities, determine research and training priorities, and allocate treatment and rehabilitation resources. In this period, the epidemiology of drug abuse has assumed great importance, and a variety of research approaches have been devised to aid in the study of heroin abuse patterns.

The most prominent and traditional of these approaches is the National Household Survey (Cisin et al. 1978) that applies standard survey research and sampling theory to the study of drug abuse prevalence in a sample of U.S. households. Because of the great cost of a National Household Survey, it cannot be implemented at (brief) regular time intervals (e.g., quarterly periods) for each of the 25 or 30 major metropolitan communities of the nation. In addition, traditional survey research methods are difficult to apply to criminal activities such as heroin abuse because of denial of the activity and the fact that the proportion of heroin users in the general population is relatively small (Hunt and Chambers 1976).

In order to supplement the national surveys with community-specific and (regular) quarterly estimates of heroin abuse prevalence, a number of researchers have adapted approaches to population size estimation used in biology and paleobiology. One such method is the so-called multiple-recapture census. In the remainder of this paper we will describe two different varieties of the multiple-recapture census and provide examples of how it can be applied to the study of heroin abuse.

THE MULTIPLE-RECAPTURE CENSUS

When applied by the biologist to study the size of animal populations, the multiple-recapture census can be described in the following way. The population under study is sampled k times. In each sample, every unmarked animal is marked uniquely; previously marked animals have their previous captures recorded; and then all animals are released back into the population. At the end of the k sampling trials, the complete capture history of every captured animal can be constructed.

Mathematical models have been formulated to estimate the size of the population from the multiple-recapture record of population members recorded on the k sampling trials. The mathematical models of this sampling process require certain technical assumptions about the nature of the sampling; for example, about the sampling probabilities for members of the population and the stability of the population during the multiple-recapture experiment. Multiple-recapture models often are classified according to their assumptions about whether the population is changing ("open") or nonchanging ("closed"). Because both the open and closed population models have potential for heroin abuse research, we present a brief description of each model, followed by applications to data from the Client Oriented Data Acquisition Process (CODAP).

Open Population Models

Open population models are useful not only to estimate population size at specific time periods, but also to estimate the number of new members entering the population between time periods and the probability of remaining in the population during a succeeding time interval (survival probability). When an extensive amount of recapture information can be constructed for the members of an open population, the Jolly-Seber model is probably of greatest potential interest to drug abuse epidemiologists (Jolly 1965; Seber 1965, 1973). When only limited recapture information is available, the more restrictive band-recovery model can be used.

The Jolly-Seber Model. In the Jolly-Seber open-population model, samples are drawn from the population at k successive time periods. At the i^{th} time period (sample) there are N_i members in the population, of which n_i were observed in the sample. To apply the Jolly-Seber model, the following recapture history must be constructed:

1) n_i = number of members in i^{th} sample

2) m_i = number of marked members in i^{th} sample

3) r_i = number of marked members released after the i^{th} sample and subsequently recaptured

4) z_i = number of members captured before the i^{th} sample that are not caught in i^{th} sample, but are caught subsequently

When certain assumptions (to be stated later) are true, then the above observed information can be used to estimate the number of members in the population at each time period, the survival probability, and the number of members entering (or leaving) the population between time periods. Computer programs now exist to compute the maximum likelihood (ML) estimates and standard errors of the above unknowns, and certain intuitively appealing estimators can be used to increase intuitive understanding of the model (or used as intermediate values in computing the ML estimates). The intuitive estimators can be explained as follows. Suppose M_i is the unknown number of marked animals in the population just prior to the i^{th} sample. Then m_i is the number of these actually captured in sample i. When certain sampling assumptions are met, the proportion of members marked in the sample (m_i/n_i) will equal the proportion of marked members in the population (M_i/N_i) and

$$\frac{m_i}{n_i} = \frac{M_i}{N_i}$$

thus

$$\hat{N}_i = \frac{M_i n_i}{m_i}.$$

Since M_i is unknown, it must be estimated from the recapture history of the members captured at each of the k samples. To estimate M_i, note that when the n_i members of the i^{th} sample are released, there are two groups of marked members. There are the marked members not captured in the i^{th} sample, $(M_i - m_i)$, of which z_i subsequently are caught, and there are n_i just released members of which r_i subsequently are caught. Under the sampling assumptions, the chances of recapture are assumed identical for both groups and

$$\frac{z_i}{M_i - m_i} = \frac{r_i}{n_i}.$$

This yields

$$\hat{M}_i = n_i(z_i/r_i) + m_i$$

as the required estimator of M_i. Then

$$\hat{N}_i = \frac{\hat{M}_i n_i}{m_i}.$$

The estimator of $\hat{M}$ also is used to obtain the estimator of ϕ_i, the survival probability, as

$$\hat{\phi}_i = \hat{M}_{i+1}/(\hat{M}_i - m_i + n_i).$$

Finally, the number of members entering the population between sample i and i+1 can be estimated by

$$\hat{\beta}_i = \hat{N}_{i+1} - \hat{\phi}_i \hat{N}_i.$$

Additional discussions of these estimators and their asymptotic variances can be found in Seber (1973).

The assumptions required for the application of the Jolly-Seber model are summarized below:

1) each member of the population has the same probability of being captured in the i^{th} sample

2) every marked member has the same probability (ϕ_i) of surviving from the i^{th} to the $(i+1)^{th}$ sample

3) every member caught has the same probability (v_i) of being returned to the population

4) marked members retain their marks

The effects of departure from the above assumptions have been well documented. Heterogeneous encounter probabilities have been shown to produce negative bias in the estimates of N_i (Carothers 1973; Gilbert 1973). When the population is heterogeneous, grouping into more homogenous subgroups (e.g., heavy vs. light users) will help alleviate the problem. Heterogeneity of survival probabilities will result in a negative bias on N_i. Several methods now exist for assessing the reasonableness of the assumptions of the model. Specific tests are available for heterogeneity of encounter probability (Leslie 1958; Carothers 1971) and heterogeneous survival probabilities (Balser 1981). Computer programs exist for computing the maximum likelihood estimates and standard errors from the open population study (Davies 1971; White 1971; Arnason and Baniuk 1980).

Closed Population Models

When the population under study is constant, then simpler multiple-recapture designs can be used to estimate the size of the closed population. These designs are popular because of their simplicity and because the estimates still may be valid even when the population changes. For example, if the population is changing between two time periods as a result of random mortality, Robson (1969) shows that certain estimators are suitable for estimating the population size at the time of the first sample. When there is an increase in the population size between samples, Seber (1973) shows that the estimate is suitable for the time of the second sample. If a high degree of temporal precision is not required (e.g., time of first sample vs. time of second sample), then these violations of the closed population assumption may present no problem. The assumption of the closed population models are:

1) the population is closed

2) all members of the population have the same probability of being captured

3) each sample is a representative sample of the population

4) the multiple-recapture history of each captured member is accurate

As in the open population models, heterogeneity of capture probabilities can be handled by grouping into homogeneous groups.

There have been a number of applications of the closed population model to the study of heroin abuse (Greenwood 1971; French 1977; Woodward et al. 1984). Most frequently a simple, two-sample design has been used that requires the additional assumption of independence among the samples. In the second example provided here, a three-sample design will be used to avoid the overly restrictive assumption of independent samples.

The three-sample multiple-recapture census for the closed population can be represented as a three-way contingency table as shown in table 1.

TABLE 1. *Schematic representation of the three-sample multiple recapture experiment*

		Captured Sample 2		Sample 3	Not captured Sample 2	
		Captured	Not captured		Captured	Not captured
Sample 1	Captured	f_{111}	f_{121}	Captured	f_{112}	f_{122}
	Not captured	f_{211}	f_{221}	Not captured	f_{212}	(f_{222})

The frequency f_{111} represents the number of individuals captured on all three samples. The other cells represent different combinations of captured and not captured members on the three sampling trials. The unobserved frequency f_{222} represents the number of persons not captured on any sampling trial.

An estimate of the total population size and its standard error can be obtained from a log-linear model where the missing cell (f_{222}) is treated as a structural zero (Bishop et al. 1975). In the three-sample capture-recapture design, the three-way interaction must be assumed to be zero; however, the samples may be pair-wise dependent.

Recently, a new approach to estimating the size of the closed population has been developed using closed-form, constrained estimators of the parameters of the log-linear model (Bonett et al., in press). This approach has several new features that will be important in drug abuse applications:

1) all estimators and standard errors are in closed form and thus can be computed easily, even in the microcomputer

2) the models can be fit under exact linear and stochastic constraints on the parameters of the log-linear model and on the unobserved population sizes, thus greatly reducing standard errors in many applications

3) the population can be divided into strata that are homogeneous in the capture-probabilities, so that the assumption of equal capture-probabilities can be met

4) hypothesis testing can be carried out on unobserved population sizes, so that differences among communities and differences across time can be studied simultaneously in a statistical framework

Application to Heroin Abuse Research

In applying the multiple-recapture census to heroin abuse, some method of sampling from the population of heroin abusers must be devised, as must a way of recording their "capture" history. Such sampling cannot be so straightforward as it is in the study of wildlife populations. In the case of heroin abuse, the sampling will be more similar to paleobiology applications where naturalistic samples of extinct animal fossils are collected from different strata of rock.

In application to heroin abuse, admissions to federally funded treatment programs during a specific time period represent a sample from the population of heroin abusers. Because the sampling is naturalistic, the population being studied consists of all heroin abusers who have a nonzero probability of entering federally funded treatment programs. This definition of the population may be more restricted than if random sampling were involved, but the population of persons who are likely to enter a federally funded treatment program is a large and important population to study.

In order to construct the multiple-capture history for heroin abusers, a computerized matching procedure was developed for the anonymous CODAP files (Woodward et al. 1984). Using 2-month periods to accumulate admissions, the computerized matching program was used to identify readmissions. A number of practical issues concerning the feasibility and accuracy of this method of constructing the multiple-recapture table are discussed elsewhere (Woodward et al. 1984).

The admission to treatment could be used as a basis for either an open or closed population model. At the local level of the treatment program, accurate records of readmission to treatment facilities could be accumulated so that an open or closed population model could be applied. In the closed population applications presented here, inaccuracies exist because a matching procedure had to be devised to recognize readmissions of anonymous clients to treatment programs, using the aggregate CODAP files.

Example 1: Homogeneity of capture probabilities. If the assumption of homogeneous capture probabilities is violated, then it is recommended that the population be stratified into groups more homogeneous in the capture probabilities. By including a relevant blocking variable in the capture-recapture design as an additional factor, the assumptions of the model may be satisfied. A variable hypothesized to be related to the probability of admission to a drug treatment program is the level of heroin use. Heroin users were identified each time they were admitted to drug treatment programs using the computerized matching procedure applied to the CODAP files. At the time of first admission, users were classified as being either heavy or moderate users of heroin. The data for a single Standard Metropolitan Statistical Area (SMSA) appear in table 2 (Doscher and Woodward 1983).

TABLE 2. *Multiple-recapture census with heavy and moderate heroin users*

	Heavy			Moderate	
	A	N		A	N
A	51	496	A	12	169
N	498	--	N	140	--

A = admitted to treatment program; N = not admitted to treatment program

This table contains two unobservable cells that correspond to the number of heavy and moderate heroin users who were not admitted during either of two admission periods. In this design the presence of two latent cells renders unestimable the Admission 1 by Admission 2 interaction effect and the Admission 1 by Admission 2 by Use interaction. An exploratory analysis revealed four constraints consistent with the data:

1) the main effect of Admission 1 equals the main effect of Admission 2

2) the Use by Admission 1 interaction effect is zero

3) the Use by Admission 2 interaction effect is zero

4) the number of moderate heroin users not admitted on either occasion is three times larger than the number of heavy heroin users not admitted on either occasion

The constrained closed-form estimators and associated test statistic of Bonett et al. (in press) yielded an excellent fit to the data (X^2 = 5.73 on 4 degrees of freedom). Given this model, it can be seen that the heavy-moderate stratification did not reveal heterogeneous capture probabilities as evidenced by the zero Use by Admission interactions. In this case, the obtained estimate of 8,152 with standard error 934 is not substantially different from the value that would have been obtained from an analysis of the design collapsing across the stratification Use factor. Although it is possible that other stratification factors may reveal heterogeneous capture probabilities, they must be discovered empirically through a study analogous to the one presented above. For the moment, the hypothesis that heterogeneous capture probabilities arise because of differences among heavy and moderate heroin users is rejected.

Example 2: A cross-sectional, longitudinal study of heroin abuse. Using the general statistical framework of the Bonett et al. constrained closed-form estimators, it is possible to carry out a larger study that addresses simultaneously differences among communities and differences across time. Twelve major metropolitan areas (SMSAs) were studied at each of three time periods. For each community, a three-sample multiple-recapture census (table 1) was conducted at each of 3 years--1977, 1978, and 1979--using the computerized matching procedure applied to the CODAP files. The 12 communities were grouped into 3 heroin supply-source clusters: 1) Mexican; 2) Southeast/Southwest Asian[1], and 3) mixed (both of the above). The grouping is, of course, approximate and represents current knowledge of supply derived from extensive drug enforcement activities. The complete design for the multiple-community multiple-time study is represented schematically in table 3.

[1]These sources include Pakistan, Afghanistan, and the Golden Triangle.

TABLE 3. *Schematic representation of the cross-sectional, longitudinal multiple-recapture study*

Supply Source	SMSA	1977	1978	1979
Mexican	City 1			
	City 2			
	City 3			
	City 4			
Southeast/ Southwest Asian	City 5			
	City 6			
	City 7			
	City 8			
Mixed	City 9			
	City 10			
	City 11			
	City 12			

Note: Within each of the 36 cells of the design, a three-sample multiple-recapture census like that of table 1 was carried out.

Within each of the 36 cells of the design was a 3-sample multiple-recapture table like the one in table 1. Now a single statistical model was fit to the entire design, so that there is a simultaneous estimation of the population sizes for all 12 communities at all 3 time periods. This provides for a statistical framework within which hypotheses can be tested about differences across communities, across time, and about the interaction of communities and time. The simultaneous estimation also permits constraints to be imposed across similar communities and across time in order to reduce the standard errors of the resulting population size estimates. This is an important contribution of the Bonett et al. methodology since a typical multiple-recapture estimation procedure (Bishop et al. 1975) may yield extremely large standard errors. By constraining certain parameters to be equal across time, and across similar communities, the standard errors were reduced substantially in the simultaneous analysis of the data.

Without such a procedure, many standard errors are as large as the estimates themselves, indicating an unacceptable degree of precision in the single-community single-time multiple-recapture census approach.

The simultaneous population size estimates and their standard errors are shown in table 4. Since the purpose of these analyses is to illustrate the methodology and to compare time trends in the different supply-source groupings, the SMSAs are referred to by arbitrary numbers. Even though a very substantial reduction in standard error size was achieved with the current methodology, some of the standard errors still are quite large. Using the estimates and their standard errors in table 4, 95 percent confidence intervals were computed for each community at each year and are presented in table 5. As can be seen, in certain cases, the precision of the multiple-recapture census is low.

TABLE 4. *Capture-recapture heroin abuse estimates and standard errors (in parentheses)*

Source	City	1977	1978	1979
Mexican	1	66,077 (10,909)	39,678 (6,582)	25,601 (6,297)
	2	5,628 (907)	4,599 (968)	12,226 (2,955)
	3	20,043 (7,527)	11,475 (4,017)	18,770 (3,681)
	4	30,404 (7,163)	44,269 (7,481)	28,396 (4,834)
Southeast/ Southwest Asian	5	48,745 (10,662)	15,863 (3,601)	15,646 (3,590)
	6	14,579 (3,251)	46,586 (7,476)	41,071 (6,827)
	7	65,044 (14,495)	4,235 (564)	29,289 (5,710)
	8	104,601 (44,837)	40,748 (10,961)	50,443 (8,728)
Mixed	9	20,302 (6,207)	17,321 (5,220)	12,948 (4,152)
	10	43,351 (6,186)	22,539 (3,313)	40,062 (6,466)
	11	63,636 (7,926)	55,672 (13,138)	27,562 (4,345)
	12	27,374 (6,651)	49,018 (12,181)	29,766 (7,334)

TABLE 5. *95 percent confidence intervals for capture-recapture heroin abuse estimates*

Source	City	1977	1978	1979
Mexican	1	44,695-87,458	26,777-52,579	13,258-37,943
	2	3,850-7,405	2,702-6,496	6,434-18,017
	3	5,290-34,795	3,602-19,348	11,555-25,984
	4	16,364-44,443	29,606-58,932	18,921-37,870
Southeast/ Southwest Asian	5	27,847-69,642	8,805-22,921	8,609-22,682
	6	8,207-20,950	31,933-61,239	27,690-54,451
	7	36,633-93,454	3,130-5,340	18,097-40,480
	8	16,720-192,481	19,264-62,232	33,336-67,549
Mixed	9	8,136-32,467	7,090-27,552	4,810-21,086
	10	31,226-55,475	16,046-29,032	27,389-52,735
	11	48,101-79,170	29,922-81,422	19,046-36,078
	12	14,338-40,409	25,143-72,893	21,271-38,261

Based on the estimates and their standard errors, it now is possible to test hypotheses about population size across communities and across time. First, we tested the (null) hypothesis that the patterns across time are equal in the population for the three source groups. This hypothesis was rejected ($\chi^2 = 20.52$, df = 4). The conclusion is that heroin abuse population sizes show different time trends for the three groups of communities (i.e., Mexican, Southeast/Southwest Asian, and mixed). The average sample values for the source time trends appear in table 6.

TABLE 6. *Average time trends in population size estimates for Mexican, Southeast/Southwest Asian, and mixed heroin supply communities*

Source	1977	1978	1979
Mexican	30,538	25,005	21,248
Southeast/ Southwest Asian	58,242	26,858	34,112
Mixed	38,665	36,137	27,584

Because the time trends differ across groups, we tested the hypothesis of zero change across time for each group. For the

Mexican-supplied communities there is a significant linear decrease across the three years, 1977, 1978, and 1979 ($X^2 = 5.12$, df = 1); for the Southeast/Southwest Asian, there is a highly significant "U" shaped pattern ($X^2 = 89.3$, df = 2). In the mixed group of communities there was a significant decrease between 1978 and 1979 ($X^2 = 6.6$, df = 2). Even though there are differences in time trends for the Mexican, Southeast/Southwest Asian, and mixed communities during 1977, 1978, and 1979, the overall pattern is one of a decrease in heroin abuse population size.

DISCUSSION

The multiple-recapture census can provide a practical method of estimating the size of a population when standard survey research methods are difficult to apply. Open population models deal with populations that are changing. The major purposes of the census are to estimate the population size at each of several time periods, the probability of remaining in the population between each time interval, and the number of persons entering or leaving the populations between each time period. Such information would be useful to the heroin abuse researcher, and practical ways of conducting the open population study of the heroin abusing population do exist.

The simpler closed population multiple-recapture census is appropriate when the population is not changing or when temporal precision is not of great importance. When the population is decreasing, for example, closed population estimates are appropriate for the time period of the first sample.

Several applications of the multiple-recapture census to heroin abuse are presented. In the first example, the issue of equal capture probabilities is examined. In the second example, a cross-sectional longitudinal design is used to test for population size differences among communities and across time. The time trends for Mexican, Southeast/Southwest Asian, and mixed supply-source cities differed in their time trends on heroin abuse population size. The overall trend across the years 1977, 1978, and 1979 was seen to be decreasing.

The applications presented here used admissions to treatment as the sampling procedure. This kind of naturalistic sampling is not inconsistent *per se* with the statistical assumptions of the multiple-recapture census, but it does restrict the definition of the population to those heroin abusers that have a nonzero probability of entering treatment. While this is a large and important population to study, this procedure does not yield what can be considered estimates of heroin abuse *prevalence* in a strict sense of the word.

One difficulty with the multiple-recapture census is the magnitude of the standard errors that result from this kind of sampling. Even the newly developed constrained closed form estimators of

Bonett et al. (in press) did not reduce all standard errors in the examples to acceptable levels. In spite of this, however, the degree of precision was sufficient to detect significant differences in population sizes across time. In addition, the time trends for communities believed to have different sources of heroin were significantly different. The overall trend during the 1977-1979 period was significantly decreasing; thus, the findings of this application are consistent with the findings of many other studies based on different sampling methodologies (DEA 1983a, 1983b; NIDA 1983).

REFERENCES

Arnason, A.N., and Baniuk, L.A. Computer system for mark-recapture analysis of open populations. J Wildl Management, 44:325-332, 1980.

Balser, J.P. Confidence interval estimation and tests for temporary outmigration in tag-recapture studies. Doctoral Dissertation, Cornell University, 1981.

Bishop, Y.M.M.; Fienberg, S.E.; and Holland, P.W. Discrete Multivariate Analysis: Theory and Practice. Cambridge, Mass.: MIT Press, 1975.

Bonett, D.G.; Woodward, J.A.; and Bentler, P.M. A linear model for estimating the size of a closed population. In press.

Carothers, A.D. An examination and extension of Leslie's test of equal catchability. Biometrics, 27:615-630, 1971.

Carothers, A.D. The effects of unequal catchability on Jolly-Seber estimates. Biometrics, 29:79-100, 1973.

Cisin, L.H.; Miller, J.D.; and Harrell, A.V. Highlights from the National Survey on Drug Abuse: 1977. Prepared for the National Institute on Drug Abuse by the Social Research Group of The George Washington Univ. and Response Analysis Corp. Washington, D.C.: Supt. of Docs., U.S. Govt. Print. Off., 1978.

Davies, R.G. Computer Programming in Quantitative Biology. London: Academic Press, 1971.

Doscher, M.L., and Woodward, J.A. Estimating the size of subpopulations of heroin users: Applications of log-linear models to capture/recapture sampling. Int J Addict, 18:167-182, 1983.

Drug Enforcement Administration. Monitor Program Report. Washington, D.C., 1983a.

Drug Enforcement Administration. Signature Report. Washington, D.C., 1983b.

French, J.F. Prevalence of Heroin Use in New Jersey. Unpublished report. Department of Health; Alcohol, Narcotic, and Drug Abuse Unit, State of New Jersey. Trenton, 1977.

Gilbert, R.O. Approximations of the bias in the Jolly-Seber capture-recapture model. Biometrics, 29:501-526, 1973.

Greenwood, J.A. Estimating Relative Incidence of Narcotic Addiction Using Treatment Facility Admissions. DEA Report RTP-TR11, June, 1974.

Hunt, L.G., and Chambers, C.D. The Heroin Epidemics. New York: Spectrum, 1976.

Jolly, G.M. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika, 52:225-247, 1965.

Leslie, P.H. Statistical appendix. J Anim Ecol, 27:84-86, 1958.

National Institute on Drug Abuse. Heroin Work Group. Unpublished report. Washington, D.C., 1983.

Robson, D.S. Mark-recapture methods of population estimation. In: Johnson, N.L., and Smith, H., Jr., eds. New Developments in Survey Sampling. New York: Wiley and Sons, 1969.

Seber, G.A.F. A note on the multiple-recapture census. Biometrika, 52:225-247, 1965.

Seber, G.A.F. The Estimation of Animal Abundance and Related Parameters. London: Griffin, 1973.

White, E.G. A Computer Program for Capture-Recapture Studies of Animal Populations: A FORTRAN Listing for the Stochastic Model of G.M. Jolly. Tussock Grasslands and Mountain Lands Institute. Spec. Publ. No. 8, 1971.

Woodward, J.A.; Retka, R.L.; and Ng, L. Construct validity of heroin abuse estimators. Int J Addict, 19:93-117, 1984

AUTHORS

J. Arthur Woodward, Ph.D.
M.L. Brecht, Ph.D.

Department of Psychology
University of California, Los Angeles
Los Angeles, CA

Douglas G. Bonett, Ph.D.
University of Southern California

PARTICIPANTS

Work Group on Validity Issues in Self-Reported Drug Use
May 8-9, 1984

Edgar Adams, M.S., Chief, SEAB
Division of Epidemiology and
Statistical Analysis, NIDA
5600 Fishers Lane, Room 11A56
Rockville, MD 20857
(301) 443-6637

Michael Backenheimer, Ph.D.
Division of Epidemiology and
Statistical Analysis, NIDA
5600 Fishers Lane, Room 11A56
Rockville, MD 20857
(301) 443-2974

Catherine Bell, M.S.
Division of Clinical Research,
NIDA
5600 Fishers Lane, Room 10A20
Rockville, MD 20857
(301) 443-1514

Marc Brodsky, M.S.
Division of Epidemiology and
Statistical Analysis, NIDA
5600 Fishers Lane, Room 11A55
Rockville, MD 20857
(301) 443-6637

Reuben Cohen, M.A.
Senior Vice-President
Response Analysis Corporation
Research Park, Route 206
Princeton, NJ 08540
(609) 921-3333

Raquel Crider, Ph.D.
Division of Epidemiology and
Statistical Analysis, NIDA
5600 Fishers Lane, Room 11A55
Rockville, MD 20857
(301) 443-6637

Blanche Frank, Ph.D., Chief
Evaluation, Research,
and Planning
New York State Division of
Substance Abuse Services
2 World Trade Center
New York, NY 10047
(212) 488-3967

Joseph Gfroerer, B.A.
Division of Epidemiology and
Statistical Analysis, NIDA
5600 Fishers Lane, Room 11A55
Rockville, MD 20857
(301) 443-6637

Bernard A. Gropper, Ph.D.
Office of Research Programs
National Institute of Justice
Washington, DC 20531
(202) 724-7631

Adele Harrell, Ph.D.
Institute for Social Analysis
1625 "K" Street
Washington, DC 20006
(202) 728-1059

Richard Hawks, Ph.D.
Chief, Research Technology
Division of Preclinical
Research, NIDA
5600 Fishers Lane, Room 10A19
Rockville, MD 20857
(301) 443-5280

Lloyd Johnston, Ph.D.
Institute for Social Research
The University of Michigan,
Box 1248
Ann Arbor, MI 48106
(313) 763-5043

Nicholas Kozel, M.S.
Division of Epidemiology and
Statistical Analysis, NIDA
5600 Fishers Lane, Room 11A40
Rockville, MD 20857
(301) 443-6637

Douglas Lipton, Ph.D.
Deputy Director for Substance
Abuse Research and
Evaluation
State of New York Division of
Substance Abuse Services
2 World Trade Center,
67th Floor
New York, NY 10047
(212) 488-3969

Jacqueline P. Ludford, M.S.
Chief, Research Analysis Branch
Office of Science, NIDA
5600 Fishers Lane
Rockville, MD 20857
(301) 443-1056

Judith Miller, Ph.D.
Social Research Group
George Washington University
2000 L Street, Northwest
Washington, DC 20036
(202) 275-1910

David Nurco, D.S.W.
Research Professor
Department of Psychiatry
University of Maryland
1229 West Mount Royal Avenue
Baltimore, MD 21217
(202) 621-4208

Louise Richards, Ph.D.
Division of Epidemiology and
Statistical Analysis, NIDA
5600 Fishers Lane, Room 11A56
Rockville, MD 20857
(301) 840-9348

Beatrice Rouse, Ph.D.
Division of Epidemiology and
Statistical Analysis, NIDA
5600 Fishers Lane, Room 11A56
Rockville, MD 20857
(301) 443-2974

Neil H. Sampson, M.B.A.,
M.P.H., Director
Division of Epidemiology and
Statistical Analysis, NIDA
5600 Fishers Lane, Room 11A56
Rockville, MD 20857
(301) 443-6504

Ray Shreckengost, M.S.,
M.A., M.P.A.
CIA
Washington, DC 20505
(703) 351-6024

Monroe Sirken, Ph.D.
National Center for
Health Statistics
Prince George's Center
Building, Room 2-12
3700 East-West Highway
Hyattsville, MD 20782
(301) 436-7110

Reginald Smart, Ph.D.
Director, Program Development
Research Department
Addiction Research Foundation
of Ontario
33 Russell Street
Toronto, Ontario
CANADA M5S 2S1
(416) 595-6017

Barry Zuckerman, M.D.
Department of Pediatrics
Boston University School
of Medicine
Boston City Hospital
818 Harrison Avenue
Boston, MA 02118
(617) 424-4234

J. Arthur Woodward, Ph.D.
Department of Psychology, UCLA
Los Angeles, CA 10024
(213) 825-6403

National Institute on Drug Abuse

Research monograph series

While limited supplies last, single copies of the monographs may be obtained free of charge from the National Clearinghouse for Drug Abuse Information (NCDAI). Please contact NCDAI also for information about availability of coming issues and other publications of the National Institute on Drug Abuse relevant to drug abuse research.

Additional copies may be purchased from the U.S. Government Printing Office (GPO) and/or the National Technical Information Service (NTIS) as indicated. NTIS prices are for paper copy. Microfiche copies, at $4.50, are also available from NTIS. Prices from either source are subject to change.

Addresses are:

NCDAI
National Clearinghouse for Drug Abuse Information
Room 10A-43
5600 Fishers Lane
Rockville, Maryland 20857

GPO
Superintendent of Documents
U.S. Government Printing Office
Washington, D.C. 20402

NTIS
National Technical Information Service
U.S. Department of Commerce
Springfield, Virginia 22161

1 FINDINGS OF DRUG ABUSE RESEARCH. Not available from NCDAI.
Vol. 1: GPO out of stock NTIS PB #272 867/AS $32.50
Vol. 2: GPO out of stock NTIS PB #272 868/AS $29.50

2 OPERATIONAL DEFINITIONS IN SOCIO-BEHAVIORAL DRUG USE RESEARCH 1975. Jack Elinson, Ph.D., and David Nurco, Ph.D., eds. Not available from NCDAI.
GPO out of stock NTIS PB #246 338/AS $16

3 AMINERGIC HYPOTHESES OF BEHAVIOR: REALITY OR CLICHE? Bruce J. Bernard, Ph.D., ed. Not available from NCDAI.
GPO Stock #017-024-00486-3 $6.50 NTIS PB #246 687/AS $16

4 NARCOTIC ANTAGONISTS: THE SEARCH FOR LONG-ACTING PREPARATIONS. Robert Willette, Ph.D., ed.
GPO out of stock NTIS PB #247 096/AS $8.50

5 YOUNG MEN AND DRUGS: A NATIONWIDE SURVEY. John A. O'Donnell, Ph.D., et al. Not available from NCDAI.
GPO Stock #017-024-00511-8 $6.50 NTIS PB #247 446/AS $16

6 EFFECTS OF LABELING THE "DRUG ABUSER": AN INQUIRY. Jay R. Williams, Ph.D. Not available from NCDAI.
GPO Stock #017-024-00512-6 $4.75 NTIS PB #249 092/AS $8.50

7 CANNABINOID ASSAYS IN HUMANS. Robert Willette, Ph.D., ed.
GPO Stock #017-024-00510-0 $6.00 NTIS PB #251 905/AS $14.50

8 Rx: 3x/WEEK LAAM - ALTERNATIVE TO METHADONE. Jack Blaine, M.D., and Pierre Renault, M.D., eds.
Not available from GPO NTIS PB #253 763/AS $14.50

9 NARCOTIC ANTAGONISTS: NALTREXONE PROGRESS REPORT. Demetrios Julius, M.D., and Pierre Renault, M.D., eds. Not available from NCDAI.
GPO Stock #017-024-00521-5 $7.00 NTIS PB #255 833/AS $17.50

10 EPIDEMIOLOGY OF DRUG ABUSE: CURRENT ISSUES. Louise G. Richards, Ph.D., and Louise B. Blevens, eds. Not available from NCDAI.
GPO Stock #017-024-00571-1 $6.50 NTIS PB #266 691/AS $22

11 DRUGS AND DRIVING. Robert Willette, Ph.D., ed. Not available from NCDAI.
GPO Stock #017-024-00576-2 $5.50 NTIS PB #269 602/AS $16

12 PSYCHODYNAMICS OF DRUG DEPENDENCE. Jack D. Blaine, M.D., and Demetrios A. Julius, M.D., eds. Not available from NCDAI.
GPO Stock #017-024-00642-4 $5.50 NTIS PB #276 084/AS $17.50

13 COCAINE: 1977. Robert C. Petersen, Ph.D., and Richard C. Stillman, M.D., eds. Not available from NCDAI.
GPO Stock #017-024-00592-4 $6.00 NTIS PB #269 175/AS $19

14 MARIHUANA RESEARCH FINDINGS: 1976. Robert C. Petersen, Ph.D., ed. Not available from NCDAI.
GPO out of stock NTIS PB #271 279/AS $22

15 REVIEW OF INHALANTS: EUPHORIA TO DYSFUNCTION. Charles Wm. Sharp, Ph.D., and Mary Lee Brehm, Ph.D., eds.
GPO Stock #017-024-00650-5 $7.50 NTIS PB #275 798/AS $28

16 THE EPIDEMIOLOGY OF HEROIN AND OTHER NARCOTICS. Joan Dunne Rittenhouse, Ph.D., ed. Not available from NCDAI.
GPO Stock #017-024-00690-4 $6.50 NTIS PB #276 357/AS $20.50

17 RESEARCH ON SMOKING BEHAVIOR. Murray E. Jarvik, M.D., Ph.D., et al., eds. Includes epidemiology, etiology, consequences of use, and approaches to behavioral change. From a NIDA-supported UCLA conference.
GPO Stock #017-024-00694-7 $7.50 NTIS PB #276 353/AS $29.50

18 BEHAVIORAL TOLERANCE: RESEARCH AND TREATMENT IMPLICATIONS. Norman A. Krasnegor, Ph.D., ed. Theoretical and empirical studies of nonpharmacologic factors in development of drug tolerance.
GPO Stock #017-024-00699-8 $5.50 NTIS PB #276 337/AS $16

19 THE INTERNATIONAL CHALLENGE OF DRUG ABUSE. Robert C. Petersen, Ph.D., ed. Papers from the VI World Congress of Psychiatry.
GPO Stock #017-024-00822-2 $7.50 NTIS PB #293 807/AS $28

20 SELF-ADMINISTRATION OF ABUSED SUBSTANCES: METHODS FOR STUDY. Norman A. Krasnegor, Ph.D., ed. Techniques used to study basic processes underlying abuse of drugs, ethanol, food, and tobacco.
GPO Stock #017-024-00794-3 $6.50 NTIS PB #288 471/AS $22

21 PHENCYCLIDINE (PCP) ABUSE: AN APPRAISAL. Robert C. Petersen, Ph.D., and Richard C. Stillman, M.D., eds. For clinicians and researchers, assessing the problem of PCP abuse.
GPO Stock #017-024-00785-4 $7.00 NTIS PB #288 472/AS $25

22 QUASAR: QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIPS OF ANALGESICS, NARCOTIC ANTAGONISTS, AND HALLUCINOGENS. Gene Barnett, Ph.D.; Milan Trsic, Ph.D.; and Robert Willette, Ph.D.; eds. Not available from NCDAI.
GPO Stock #017-024-00786-2 $8.00 NTIS PB #292 265/AS $35.50

23 CIGARETTE SMOKING AS A DEPENDENCE PROCESS. Norman A. Krasnegor, Ph.D., ed. Discusses factors involved in the onset, maintenance, and cessation of the cigarette smoking habit. Includes an agenda for future research.
GPO Stock #017-024-00895-8 $6.00 NTIS PB #297 721/AS $19

24 SYNTHETIC ESTIMATES FOR SMALL AREAS: STATISTICAL WORKSHOP PAPERS AND DISCUSSION. Jos. Steinberg, ed. Papers from a workshop on statistical approaches that yield needed estimates of data for States and local areas. Not available from NCDAI.
GPO Stock #017-024-00911-3 $8.00 NTIS PB #299 009/AS $23.50

25 BEHAVIORAL ANALYSIS AND TREATMENT OF SUBSTANCE ABUSE. Norman A. Krasnegor, Ph.D., ed. Papers on commonalities and implications for treatment of dependency on drugs, ethanol, food, and tobacco.
GPO Stock #017-024-00939-3 $5.00 NTIS PB #80-112428 $22

26 THE BEHAVIORAL ASPECTS OF SMOKING. Norman A. Krasnegor, Ph.D., ed. Reprint of the behavioral section of the 1979 Report of the Surgeon General on Smoking and Health; introduction by editor.
GPO out of stock NTIS PB #80-118755 $17.50

27 PROBLEMS OF DRUG DEPENDENCE, 1979: PROCEEDINGS OF THE 41ST ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. L.S. Harris, Ph.D., ed. Not available from NCDAI.
GPO Stock #017-024-00981-4 $9.00 NTIS PB #80-175482 $37

28 NARCOTIC ANTAGONISTS: NALTREXONE PHARMACOCHEMISTRY AND SUSTAINED-RELEASE PREPARATIONS. Robert Willette, Ph.D., and Gene Barnett, Ph.D., eds. Papers report research on sustained-release and long-acting devices for use with the narcotic antagonist naltrexone. Not available from NCDAI.
GPO Stock #017-024-01081-2 $7.00 NTIS PB #81-238875 $23.50

29 DRUG ABUSE DEATHS IN NINE CITIES: A SURVEY REPORT. Louis A. Gottschalk, M.D., et al. Not available from NCDAI.
GPO Stock #017-024-00982-2 $6.50 NTIS PB #80-178882 $17.50

30 THEORIES ON DRUG ABUSE: SELECTED CONTEMPORARY PERSPECTIVES. Dan J. Lettieri, Ph.D.; Mollie Sayers; and Helen Wallenstein Pearson, eds. Volume presents summaries of major contemporary theories of drug abuse by each of 43 leading theorists.
GPO Stock #017-024-00997-1 $10.00 Not available from NTIS

31 MARIJUANA RESEARCH FINDINGS: 1980. Robert C. Petersen, Ph.D., ed. The text of the 8th Marijuana and Health report to Congress and the background scientific papers on which it was based.
GPO out of stock NTIS PB #80-215171 $20.50

32 GC/MS ASSAYS FOR ABUSED DRUGS IN BODY FLUIDS. Rodger L. Foltz, Ph.D.; Allison F. Fentiman, Jr., Ph.D.; and Ruth B. Foltz. A collection of methods for quantitative analysis of several important drugs of abuse by gas chromatography- mass spectrometry.
GPO Stock #017-024-01015-4 $6.00 NTIS PB #81-133746 $19

33 BENZODIAZEPINES: A REVIEW OF RESEARCH RESULTS, 1980. Stephen I. Szara, M.D., D.Sc., and Jacqueline P. Ludford, M.S., eds. A RAUS (Research Analysis and Utilization System) Review Report on the abuse liability of the benzodiazepine "tranquilizers."
GPO Stock #017-024-01108-8 $5.00 NTIS PB #82-139106 $13

34 PROBLEMS OF DRUG DEPENDENCE, 1980: PROCEEDINGS OF THE 42ND ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D. ed. Not available from NCDAI.
GPO Stock #017-024-01061-8 $8.00 NTIS PB #81-194847 $34

35 DEMOGRAPHIC TRENDS AND DRUG ABUSE, 1980-1995. Louise G. Richards, Ph.D., ed. Estimates of probable extent and nature of nonmedical drug use, 1980-1995, based on age structure and other characteristics of U.S. population.
GPO Stock #017-024-01087-1 $4.50. NTIS PB #82-103417 $13

36 NEW APPROACHES TO TREATMENT OF CHRONIC PAIN: A REVIEW OF MULTIDISCIPLINARY PAIN CLINICS AND PAIN CENTERS. Lorenz K.Y. Ng, M.D., ed. Discussions by active practitioners in the treatment of pain.
GPO Stock #017-024-01082-1 $5.50. NTIS PB #81-240913 $19

37 BEHAVIORAL PHARMACOLOGY OF HUMAN DRUG DEPENDENCE. Travis Thompson, Ph.D., and Chris E. Johanson, Ph.D., eds. Presents a growing body of data, systematically derived, on the behavioral mechanisms involved in use and abuse of drugs.
GPO Stock #017-024-01109-6 $6.50 NTIS PB #82-136961 $25

38 DRUG ABUSE AND THE AMERICAN ADOLESCENT. Dan J. Lettieri, Ph.D., and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report, emphasizing use of marijuana: epidemiology, sociodemographic and personality factors, family and peer influence, delinquency, and biomedical consequences.
GPO Stock #017-024-01107-0 $4.50 NTIS PB #82-148198 $14.50

39 YOUNG MEN AND DRUGS IN MANHATTAN: A CAUSAL ANALYSIS. Richard R. Clayton, Ph.D., and Harwin L. Voss, Ph.D. Examines the etiology and natural history of drug use, with special focus on heroin. Includes a Lifetime Drug Use Index.
GPO Stock #017-024-01097-9 $5.50 NTIS PB #82-147372 $19

40 ADOLESCENT MARIJUANA ABUSERS AND THEIR FAMILIES. Herbert Hendin, M.D., Ann Pollinger, Ph.D., Richard Ulman, Ph.D., and Arthur Carr, Ph.D. A psychodynamic study of adolescents involved in heavy marijuana use, to determine what interaction between family and adolescent gives rise to drug abuse.
GPO Stock #017-024-01098-7 $4.50 NTIS PB #82-133117 $13

41 PROBLEMS OF DRUG DEPENDENCE, 1981: PROCEEDINGS OF THE 43RD ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed. Not available from NCDAI.
Not available from GPO NTIS PB #82-190760 $41.50

42 THE ANALYSIS OF CANNABINOIDS IN BIOLOGICAL FLUIDS. Richard L. Hawks, Ph.D., ed. Varied approaches to sensitive, reliable, and accessible quantitative assays for the chemical constitutents of marijuana, for researchers. Not available from NCDAI.
GPO Stock #017-024-01151-7 $5 NTIS PB #83-136044 $1643

43 PROBLEMS OF DRUG DEPENDENCE, 1982: PROCEEDINGS OF THE 44TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed. Not available from NCDAI.
GPO Stock #017-024-01162-2 $8.50 NTIS PB #83-252-692/AS $40

44 MARIJUANA EFFECTS ON THE ENDOCRINE AND REPRODUCTIVE SYSTEMS. Monique C. Braude, Ph.D., and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report of animal studies and preclinical and clinical studies of effects of cannabinoids on human endocrine and reproductive functions.
GPO Stock #017-024-01202-5 $4. NTIS PB #85-150563/AS $14.50

45 CONTEMPORARY RESEARCH IN PAIN AND ANALGESIA, 1983. Roger M. Brown, Ph.D.; Theodore M. Pinkert, M.D., J.D.; and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report on the anatomy, physiology, and neurochemistry of pain and its management.
GPO Stock #017-024-01191-6 $2.75 NTIS PB #84-184670/AS $11.50

46 BEHAVIORAL INTERVENTION TECHNIQUES IN DRUG ABUSE TREATMENT. John Grabowski, Ph.D.; Maxine L. Stitzer, Ph.D., and Jack E. Henningfield, Ph.D., eds. Reports on behavioral contingency management procedures used in research/treatment environments.
GPO Stock #017-024-01192-4 $4.25 NTIS PB #84-184688/AS $16

47 PREVENTING ADOLESCENT DRUG ABUSE: INTERVENTION STRATEGIES. Thomas J. Glynn, Ph.D.; Carl G. Leukefeld, D.S.W.; and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report on a variety of approaches to prevention of adolescent drug abuse, how they can be applied, their chances for success, and needed future research.
GPO Stock #017-024-01180-1 $5.50 NTIS PB #85-159663/AS $22

48 MEASUREMENT IN THE ANALYSIS AND TREATMENT OF SMOKING BEHAVIOR. John Grabowski, Ph.D., and Catherine S. Bell, M.S., eds. Based upon a meeting cosponsored by NIDA and the National Cancer Institute to delineate necessary and sufficient measures for analysis of smoking behavior in research and treatment ettings.
GPO Stock #017-024-01181-9 $4.50 NTIS PB 84-145-184 $14.50

49 PROBLEMS OF DRUG DEPENDENCE, 1983: PROCEEDINGS OF THE 45TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed. A collection of papers which together record a year's advances in drug abuse research; also includes reports on tests of new compounds for efficacy and dependence liability.
GPO Stock #017-024-01198-3 $12 NTIS PB 85-159663/AS $22.

50 COCAINE: PHARMACOLOGY, EFFECTS, AND TREATMENT OF ABUSE. John Grabowski, Ph.D., ed. Content ranges from an introductory overview through neuropharmacology, pharmacology, animal and human behavioral pharmacology, patterns of use in the natural environment of cocaine users, treatment, through commentary on societal perceptions of use.
GPO Stock #017-020-01214-9 $4 NTIS PB 85-150381/AS $14.50

51 DRUG ABUSE TREATMENT EVALUATION: STRATEGIES, PROGRESS, AND PROSPECTS. Frank M. Tims, Ph.D., ed. A state-of-the-art review of drug abuse treatment evaluation, identifying research needs, promising approaches, and emerging issues.
GPO Stock #017-020-01218-1 $4.50 NTIS PB 85-150365/AS $17.50

52 TESTING DRUGS FOR PHYSICAL DEPENDENCE POTENTIAL AND ABUSE LIABILITY. Joseph V. Brady, Ph.D., and Scott E. Lukas, Ph.D., eds. Describes animal and human test procedures for assessing dependence potential and abuse liability of opioids, stimulants, depressants, hallucinogens, cannabinoids, and dissociative anesthetics.
GPO Stock #017-024-0204-1 $4.25 NTIS PB 85-150373/AS $16

54 MECHANISMS OF TOLERANCE AND DEPENDENCE. Charles Wm. Sharp, Ph.D., ed. Review of basic knowledge concerning the mechanism of action of opiates and other drugs in producing tolerance and/or dependence.
GPO Stock #017-024-01213-1 $8.50 NTIS PB No. to be assigned.

55 PROBLEMS OF DRUG DEPENDENCE, 1984. PROCEEDINGS OF THE 46TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.

IN PRESS OR PREPARATION

53 PHARMACOLOGICAL ADJUNCTS IN SMOKING CESSATION. John Grabowski, Ph.D., ed.

56 ETIOLOGY OF DRUG ABUSE: IMPLICATIONS FOR PREVENTION. Coryl LaRue Jones, Ph.D., and Robert J. Battjes, D.S.W., eds.

58 PROGRESS IN THE DEVELOPMENT OF COST-EFFECTIVE TREATMENT FOR DRUG ABUSERS. Rebecca S. Ashery, D.S.W., ed.

59 CURRENT RESEARCH ON THE CONSEQUENCES OF MATERNAL DRUG ABUSE. Theodore M. Pinkert, M.D., J.D., ed.

60 PRENATAL DRUG EXPOSURE: KINETICS AND DYNAMICS. C. Nora Chiang, Ph.D., and Charles C. Lee, Ph.D., eds.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1985—478-257