

TECHNICAL PAPERS ON HEALTH AND BEHAVIOR MEASUREMENT

TECHNICAL PAPER 20

Quality of Treatment Data: Reliability over Time of Self-Reports Given by Clients in Treatment for Substance Abuse

Elissa B.G. Adair, S. Gail Craddock, Heather G. Miller, and Charles F. Turner

Reference Citation

Adair, E.B.G., S.G. Craddock, H.G. Miller, and C.F. Turner. (1996) Quality of treatment data: Reliability over time of self-reports given by clients in treatment for substance abuse. *Journal of Substance Abuse Treatment* 13(2):145-149.

BRIEF REPORT

Quality of Treatment Data

Reliability Over Time of Self-Reports Given by Clients in Treatment for Substance Abuse

ELISSA B.G. ADAIR, S. GAIL CRADDOCK, HEATHER G. MILLER AND CHARLES F. TURNER

Research Triangle Institute

Abstract— *This study examines the reliability over a 2-month period of self-reports of drug use, sexual behaviors, and use of treatment services provided by 2,968 clients participating in a large, multisite, prospective study of drug treatment in the United States—the Drug Abuse Treatment Outcome Study (DATOS). Analyses focus on responses to 62 pairs of logically related questions that were asked at two points in time: (1) 1 month after entry into treatment, and (2) 3 months after entry into treatment. Subjects' responses to questions asked at these two time points are assessed for logical consistency. Prior analyses of self-reports provided by DATOS clients at one point in time (entry into treatment) found surprisingly high levels of within-interview consistency in their reporting of alcohol use (Turner & Hubbard, 1995) and cocaine use (Adair, Craddock, Miller, & Turner, 1995). The crosstemporal tests of consistency reported in this article eliminate several potential sources of artifactual consistency that may have affected prior analyses, (e.g., consistency imposed by an interviewer or constructed by a respondent during the course of a single interview). Contrary to expectations, crosstemporal comparisons reveal high levels of logical consistency in clients' responses. The mean percent of substantively inconsistent responses ranges from 0.7% for questions asking about frequency of drug use to 4.4% for questions asking about sexual behaviors.*

Keywords— self-reports; drug abuse treatment; measurement reliability; data quality.

INTRODUCTION

THERE ARE A VARIETY of reasons for both researchers and clinicians to elicit information from individuals undergoing treatment for substance abuse (Babor, 1993), such as evaluating treatment services (Hubbard, Mars-

den, Rachal, Harwood, Cavanaugh, & Ginzburg, 1989), testing hypotheses about the best patient-program fit (McLellan & Alterman, 1991), and assessing the need for additional services, such as AIDS prevention programs (Brown, 1991). There remain, however, significant questions concerning the accuracy of self-reported data in general and data on drug use in particular (Ball, 1967).

In order to assess the quality of survey data collected from individuals undergoing treatment for drug use, we have previously conducted a series of investigations using cross-sectional data from the Drug Abuse Treatment Outcome Study (DATOS), a large, multisite study conducted in the United States. These investigations have

The research reported in this article benefited greatly from conversations with Rose Etheridge, Robert Hubbard, and other colleagues at Research Triangle Institute's Center for Community Based Treatment Research Methods. This research was supported by Grant P50 DA06990-02 from the National Institute on Drug Abuse which supports the Center.

Requests for reprints should be addressed to Dr. Charles F. Turner, Director, Program in Health and Behavior Measurement, Research Triangle Institute, 6101 Executive Boulevard, Rockville, MD 20852-3909.

assessed the quality of reports of alcohol use (Turner & Hubbard, 1995) and cocaine use (Adair et al., 1995) provided by subjects at intake into DATOS. Contrary to expectations, both studies found few logically inconsistent responses. This was in marked contrast to findings from parallel analyses of responses given to similar questions by representative national samples of adults and adolescents in the National Household Survey of Drug Abuse (Cox, Witt, Traccarella, & Perez-Michael, 1992).

In considering potential explanations for the unexpectedly high levels of internal consistency found in the DATOS intake data, we speculated that two mechanisms might have artificially inflated the observed rates of consistent reporting in DATOS. First, we thought it might be possible that research interviewers imposed consistency by querying and reconciling inconsistencies they observed during the course of the interview. Secondly, it was suggested to us that most clients in treatment for drug abuse already had considerable experience replying to questions about their past drug use, and, thus, their responses would be well rehearsed.

To test these speculations, we have undertaken additional longitudinal analyses that compare reports given after an interval of 2 months had elapsed. The questions we analyze include numerous questions concerning recent behaviors; responses to these questions are less likely to be "well rehearsed." Our new analyses provide a more stringent test of the quality of these data because interviewers are unlikely to remember responses given by clients 2 months earlier and any "rehearsal" effect will be minimized for reports of recent behaviors.

STUDY DESIGN

DATOS is an ongoing prospective study of drug treatment effectiveness; it began in November, 1991. Study sites include traditional, publicly funded outpatient methadone, residential and outpatient drug-free modalities, as well as new treatment strategies that include rehabilitation services. Individuals are interviewed upon entry into treatment and at regular intervals thereafter (1, 3, 6, and 12 months). The DATOS questionnaire was designed to maximize consistency and comparability with other major research efforts (Horton, 1993). Data are collected on a range of issues, including recent drug use, mental and physical health status, social functioning, sexual behavior, involvement in illegal activities, and in-treatment experiences. (A detailed description of the DATOS protocol and sample has been published in Turner & Hubbard, in press.)

Sample

The sample for this study is comprised of DATOS clients who were in treatment long enough to complete the 3-month in-treatment interview. The present study uses data from the 2,968 respondents who were interviewed at

both 1 month (T_1) and 3 months (T_2) after intake.¹ Approximately two-thirds (66%) of respondents in this study are male. The majority of subjects are less than 40 years of age, with 35% in their 20s and 57% in their 30s. Slightly more than one-third (37%) left high school prior to graduation, and approximately the same proportion (39%) received a high school diploma. DATOS subjects are racially and ethnically diverse; 41% are white, 42% African American, and 15% Hispanic. Our sample is drawn from all four DATOS treatment modalities: residential (40%), short-term in-patient (5%), methadone (28%), and outpatient (27%). Over half (54%) of the sample was involved with the criminal justice system at the time of intake. Half (51%) had used cocaine weekly or more often in the year before treatment, while approximately one-third (34%) had used heroin at least once a week.

Analysis Strategy

This study examines responses to 33 pairs (T_1 and T_2) of questions on drug use, 24 pairs of questions on treatment received, and four pairs of questions on sexual behavior. Responses provided at 1 month after treatment are compared to those given at 3 months after treatment to assess consistency of responses. Most questions asked respondents to indicate whether or not specific events or activities had occurred recently; additional questions asked respondents to report on the frequency of some events or activities. For example, at 1 month after intake, respondents were asked, "On how many days have you been drunk since your admission?" After 3 months, respondents were asked, "On how many days have you been drunk in the last 3 months?" For responses to be logically consistent, the number of days drunk between intake and the 1-month interview must be less than or equal to the number of days drunk between intake and the 3-month interview.

Variation in consistency of responses over time may be a function of the format in which questions were asked and the response categories provided. Therefore, we conducted a separate analysis of response consistency by question format and the structure of response categories. Questions were grouped by the following categories: (1) closed-ended questions requiring only a yes or no response; (2) closed-ended questions with fixed response categories; and (3) open-ended questions that only ask respondents for a number.

In addition to assessing the consistency of responses provided by DATOS subjects, we also examined item nonresponse and survey administration errors, such as

¹This sample of respondents completing 1-month and 3-month follow-up interviews is not identical to that reported by Turner and Hubbard (1995) and Adair et al. (1995). The latter papers report analyses of data from intake interviews, and they restricted their sample to clients who completed their intake interviews after December 31, 1992.

TABLE 1
Consistency of Paired Responses (T₁-T₂) by Survey Topic

Survey Topic	No. of Q. Pairs	Substantive Inconsistency in T ₁ and T ₂ Responses	Inconsistent Skip Implying Logical Inconsistency	Inconsistent Skip but No Logical Inconsistency	T ₁ and T ₂ Not Comparable Due to Error in Survey Administration§
Drug use	14	1.9	Not Applicable. Only one question had a skip.		0.5
Freq. of drug use	5	0.7	3.4	4.8	0.7
Began in-treatment drug use in first week	7	1.0	2.2	1.7	0.8
Began in-treatment drug use after first week	7	0.9	4.1	NA	0.5
Sexual behavior‡	4*	4.4	2.5	8.6	4.2
Treatment services‡	24†	3.8	4.7	6.7	1.4
Overall	62	2.6	3.8	4.8	1.1

Table entries are mean percent of inconsistent responses across items.

*Different questions were asked of men and women. Men skipped out of the section asked of females and vice versa. Prior to consistency analysis, male and female data were aggregated.

†Of the 24 questions, 6 involved use of skip patterns.

‡A small proportion (<0.5%) of respondents refused to answer some or all of these questions on sexual behavior. Refusals were excluded from comparisons. Similarly, 1.3% of respondents who responded "Don't Know" to a treatment question asking about whether or not they had seen and discussed a treatment plan were excluded from comparison.

§Errors in survey administration include item missing data, illegible responses, out-of-range response, etc. (Respondent refusals to answer an item at T₁ or T₂ are not included; these cases were excluded from analysis.)

missing data, illegible response, and out-of-range responses. In this study, item nonresponse does not include subjects who declined to answer specific questions. The number of DATOS subjects who refused to answer questions of interest to this study was very small (less than 0.5% for any given question); they have been excluded from analyses presented in this paper.

RESULTS

Table 1 summarizes our findings concerning the consistency of responses by survey topic, and it provides an overview of item nonresponse and survey administration error. For each substantive area of questions, the mean percent of inconsistent responses to paired questions are provided along with the mean percent of skip inconsistencies and other survey administration errors.

The data presented in Table 1 indicates a high level of consistency in clients' reports over a range of topics. The overall mean percent of inconsistent responses for all 62 pairs of questions was only 2.6%. It is somewhat surprising to note fewer inconsistencies in subjects' responses to questions on frequency of drug use (mean: 0.7%; range of 0.0 to 1.9%) and drug use during first week of treatment (mean: 1.0%; range of 0.0 to 2.7%) or after first week of treatment (mean: 0.9%; range of 0.0 to 2.4%) than to questions on more neutral topics, such as treatment services (mean: 3.8%; range of 0.0 to 13.5%) or on sensitive behaviors not related to drug treatment, such as sexual behavior (mean: 4.4%; range of 0.1 to 7.4%).

There are also relatively few problems with skip patterns associated with questions of interest in this data set.² When comparing skip pattern and survey adminis-

tration error to response inconsistencies, a similar pattern emerges. The mean percent skip pattern and survey administration error is higher for questions concerning sexual behavior and treatment services than for those concerning drug use. However, overall, there are relatively few skip or survey administration errors.

The results of analyses looking at response consistency by question format and response categories are presented in Table 2. High levels of consistency are found for all three categories of questions. However, the range of inconsistent responses is broadest for questions that could be answered yes or no (0.0-13.5%). It should be noted, however, that this category included twice as many questions as were included in the other two question categories.

While the overall results of our analyses are quite encouraging, we did identify some problematic measurements. Three questions have inconsistency rates of over 8% (range: 8.2 to 13.5%). The two questions with the highest rates on inconsistent responses were: (1) "Since your admission, have you had a check-up or have you received any scheduled individual services for medical problems other than those I have already asked about?" and (2) "Since your admission, have you attended any other scheduled talks, lectures, or films as a part of your treatment?" We believe that these questions presented

²It should be noted that not all questions involved use of skip patterns. Only one question related to drug use included a skip pattern, and, therefore, no statistic for skip error was calculated for this substantive area. In addition, only 6 of the 24 questions on treatment services included a skip pattern.

TABLE 2
Consistency of Paired Responses (T₁-T₂) by
Question Format

Question Format	No. of Question Pairs Compared	Substantive Inconsistency in T ₁ and T ₂ Responses	T ₁ and T ₂ Not Comparable Due to Error in Survey Administration*
Yes or no	38	2.9 (0-13.5)	1.1 (0.1-7.9)
Closed-ended frequency	12	0.9 (0-3.0)	1.3 (0.2-4.7)
Open-ended number	12	3.2 (0-8.2)	1.1 (0.2-4.1)

Table entries are mean percent (and ranges) of inconsistent responses across items.

*Errors in survey administration include item missing data, illegible responses, out-of-range response, etc. (Respondent refusals to answer an item at T₁ or T₂ are not included; these cases were excluded from analysis.)

problems for respondents because of poorly defined terms. Thus, several respondents asked about what would be included in "scheduled individual services for medical problems" as well as talks and lectures. The third question with a high consistency rate was: "Since your admission, how much would you say you have spent on drugs for your own nonmedical use, excluding alcohol?" This question asked for a calculation that it may have been difficult for respondents to perform. Not all drugs are purchased; some are exchanged for goods or services, and some are shared in a social interaction.

DISCUSSION

Overall, our analyses indicate that DATOS subjects provide consistent reports over a 3-month period on activities and events related to drug use, sexual behavior, and use of treatment services. High levels of response consistency are found for different types of drug use, for topics other than recent drug use, and for different question formats. The literature suggests that cognitive difficulties answering these types of questions are nontrivial; such questions often require a series of calculations or cognitive operations in order to answer each question (Forsyth, Lessler, & Hubbard, 1992; Midanik & Hines, 1991). Despite these difficulties, DATOS clients provided a stable, consistent description of recent activities, with slightly more inconsistencies in reports of sexual behavior and use of treatment services.

Recent assessment of the National Household Survey on Drug Abuse (NHSDA) found much higher levels of inconsistent reporting and item nonresponse (Turner, Lessler, & Gfroerer, 1992). For instance, a comparison of questions on use or nonuse of cocaine in the past 30 days, past 12 months, and lifetime found inconsistent re-

sponses ranging from 0.18 to 22.7% (Cox et al., 1992). Thus, our results are somewhat surprising, because many of the questions in the DATOS instrument were very similar to those asked in the NHSDA, and the results we report here are for consistency in reporting over a 2-month time period. (The NHSDA does not have a longitudinal component).

One might expect that the population served by drug treatment facilities would be less adept at responding to survey questions than a national probability sample. It is not unreasonable to expect that educational achievement is lower among individuals seeking drug treatment than the general population, and there is also the possibility of drug-related impairment of cognitive functions. However, this study finds greater response consistency among DATOS subjects than that obtained in the NHSDA. Several factors may account for this unanticipated finding. It is possible that clients of drug abuse treatment centers may be more conscious of and/or more truthful about their behaviors than respondents to national surveys. Perhaps the questions posed in these studies are more appropriate for populations with a high prevalence of drug use. As Svanum and Ehrmann (1993) reported in a recent evaluation of the MacAndrew Alcoholism Scale, certain measures are less suited for use outside the clinical settings for which they were originally designed. Alternatively, patterns of inconsistent response may have been minimized in DATOS because the instrument was administered by interviewers who could verify and correct instances of (intra-interview) inconsistency. The NHSDA depends heavily on respondents to complete paper-and-pencil forms.

While our analyses do not provide a clear picture of the cause(s), our major conclusion appears inescapable after a series of both intrainterview analyses (Adair et al., 1995; Turner & Hubbard, 1995) and longitudinal analyses of measurement reliability. The self report data provided by DATOS clients during treatment for substance abuse evidences unexpectedly high levels of internal consistency—both within a single interview and across interviews conducted 2 months apart.

REFERENCES

- Adair, E.B.G., Craddock, S.G., Miller, H.G., & Turner, C.F. (1995). Assessing consistency of responses to questions on cocaine use. *Addiction*, *90*, 1497-1502.
- Babor, T.F. (1993). Alcohol and drug use history, patterns and problems. In B.J. Rounsaville, F.M. Times, A.M. Horton, & B.J. Sower (Eds.), *Diagnostic source book on drug abuse research and treatment* (NIH Publication No. 93-3508, pp. 87-92). Washington, DC: U.S. Department of Health and Human Services.
- Ball, J.C. (1967). The reliability and validity of interview data obtained from 59 narcotic drug addicts. *American Journal of Sociology*, *72*, 650-654.
- Brown, L.S. (1991). The impact of AIDS on drug abuse treatment. In R.W. Pickens, C.G. Leukefeld, & C.R. Schuster (Eds.), *Improving drug abuse treatment* (pp. 385-393). Washington, DC: NIDA, DHHS Publication No. (ADM) 91-1754.
- Cox, B.G., Witt, M.B., Traccarella, M.A., & Perez-Michael, A.M.

- (1992). Inconsistent reporting of drug use in 1988. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (pp. 109-153). Washington, DC: U.S. Department of Health and Human Services DHHS Publication Number (ADM) 92-1929.
- Forsyth, B., Lessler, J., & Hubbard, M. (1992). Cognitive evaluation of the questionnaire. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (pp. 13-52). Washington, DC: U.S. Department of Health and Human Services DHHS Publication Number (ADM) 92-1929.
- Horton, A.M. (1993). Future direction in the development of addiction assessment instruments. In B.J. Rounsaville, F.M. Times, A.M. Horton, & B.J. Sowder (Eds.), *Diagnostic source book on drug abuse research and treatment* (pp. 87-92). Washington, DC: U.S. Department of Health and Human Services, NIH Publication No. 93-3508.
- Hubbard, R.L., Marsden, M.E., Rachal, J.V., Harwood, H.J., Cavanaugh, E.R., & Ginzburg, H.M. (1989). *Drug abuse treatment: A national study of effectiveness*. Chapel Hill: University of North Carolina Press.
- McLellan, A.T., & Alterman, A.I. (1991). Patient treatment matching: A conceptual and methodological review with suggestions for future research. In R.W. Pickens, C.G. Leukefeld, & C.R. Schuster (Eds.), *Improving drug abuse treatment* (pp. 114-135). Washington, DC: NIDA, DHHS Publication No. (ADM) 91-1754.
- Midanik, L.T., & Hines, A.M. (1991). "Unstandard" ways of answering standard questions: Protocol analysis in alcohol survey research. *Drug and Alcohol Dependence*, *27*, 245-252.
- Svanum, S., & Ehrmann, L.C. (1993). The validity of the MMPI in identifying alcoholics in a university setting. *Journal of Studies in Alcohol*, *54*, 722-729.
- Turner, C.F., & Hubbard, R.L. (1995). Quality of alcohol use histories collected at intake to substance abuse treatment. *International Journal of the Addictions*, *30*(8), 963-989.
- Turner, C.F., Lessler, J.T., & Gfroerer, J.C. (Eds.). (1992). *Survey measurement of drug use: Methodological studies*. Washington, DC: U.S. Department of Health and Human Services, DHHS Publication No. (ADM) 92-1929.