Published in final edited form as:
Psychol Addict Behav. 2008 March ; 22(1): 78–87. doi:10.1037/0893-164X.22.1.78.

Factors Affecting Agreement Between Severely Mentally Ill Alcohol Abusers' and Collaterals' Reports of Alcohol and Other Substance Use

Paul R. Stasiewicz, **Paula C. Vincent**, **Clara M. Bradizza**, **Gerard J. Connors**, and **Nicole D. Mercer**
Research Institute on Addictions, University at Buffalo, State University of New York

Stephen A. Maisto
Department of Psychology, Syracuse University

Abstract

This study examined subject–collateral reports of alcohol use among a sample of 167 dually diagnosed individuals seeking outpatient treatment at a community mental health clinic. All subjects met Diagnostic and Statistical Manual of Mental Disorders (4th ed.; American Psychiatric Association, 1994) criteria for a schizophrenia-spectrum or bipolar disorder and for alcohol abuse or dependence. Subjects were recruited within 2 weeks of treatment entry and completed measures of cognitive functioning, alcohol dependence severity, psychiatric symptoms, and quantity and frequency of substance use over the previous 60 days using the Timeline Follow-Back interview (L. C. Sobell & M. B. Sobell, 1996). They also provided a urine sample, which was screened for recent drug use. Collateral interviews were conducted by phone and included an assessment of the subject's alcohol and drug use over the same 60-day period. Collaterals also reported their confidence in the accuracy of their reports. Overall, the results indicated generally poor subject–collateral agreement. However, subject–collateral agreement appeared better for those individuals (n = 97) with negative urine drug screens. The most consistent predictor of subject–collateral discrepancy scores was subjects' recent drug use. Recommendations for enhancing the validity of self-reports of substance use in a severely mentally ill population are discussed.

Keywords

alcohol; bipolar; collaterals; self-report; schizophrenia

Verbal self-report continues to be the primary method by which clinicians and researchers obtain measurements of a person's past substance use (Connors & Maisto, 2003). Collateral reports are viewed as an important second measure of an individual's drinking and drug use behavior. In a recent review of the subject–collateral (S-C) literature, Connors and Maisto (2003) concluded that “recent research supports the assertion that subjects provide accurate

Correspondence concerning this article should be addressed to Paul R. Stasiewicz, Research Institute on Addictions, 1021 Main Street, Buffalo, NY 14203. stasiewi@ria.buffalo.edu.

reports about their drinking and associated consequences" (p. 21). This conclusion was drawn from research examining S-C agreement among various clinical (e.g., outpatient alcoholics) and non-clinical (e.g., general population) subject populations.

Although information obtained from alcohol-dependent men and women tends to be reliable and valid, there can be considerable variability in accuracy depending on many factors (Babor & Del Boca, 1992). In an effort to provide a framework for understanding the optimal conditions for obtaining self-report information, Babor, Brown, and Del Boca (1990; Babor & Del Boca, 1992; Babor, Stephens, & Marlatt, 1987) proposed a conceptual model that specifies the major sources of unreliability and invalidity in the question-answering process. In brief, within the social context of the interview, the respondent's self-report is said to be influenced directly by a combination of task variables, respondent characteristics, motivation, and cognitive processes. A key benefit of this model is that it suggests ways to enhance response accuracy. In the alcohol field, most research on response accuracy has focused on respondent characteristics, such as the individual's severity of alcohol dependence and level of cognitive impairment. Several studies have demonstrated that severity of alcohol dependence is inversely related to the reliability of self-reports of drinking (e.g., Stasiewicz, Bradizza, & Connors, 1997; Toneatto, Sobell, & Sobell, 1992), whereas results from studies investigating cognitive functioning have been mixed (e.g., Babor, Steinberg, Anton, & Del Boca, 2000; Miller & Barasch, 1985).

Although considerable confidence can be placed in the accuracy of a person's self-report of drinking and in the use of collateral reports as a second measure of drinking, concerns about the accuracy of self-reports have been raised for persons with a severe mental illness (SMI; e.g., Carey & Simons, 2000), such as schizophrenia or bipolar disorder. For example, use of psychotropic medications, psychiatric symptom severity, and degree of cognitive impairment have been hypothesized to affect the accuracy of self-reports. In addition, some persons with SMI may not have reliable collateral informants, perhaps because of social isolation and/or estrangement from family (Carey & Correia, 1998). In the absence of a collateral informant, it is important to know the conditions necessary to obtain accurate self-report information from SMI patients. Because even small amounts of substance use have been associated with negative outcomes in psychiatric patients (e.g., Drake, Osher, & Wallach, 1989), it is important to evaluate the accuracy of self-reported substance use in this population.

Among studies examining the accuracy of self-reports in SMI substance abusers, only three studies have used information from collateral informants (Carey & Simons, 2000; Drake et al., 1990; Weiss, Greenfield, Griffin, Najavits, & Fucito, 2000). Carey and Simons (2000) examined the utility of collateral informants for validating self-reported substance use by 92 psychiatric outpatients at a public psychiatric hospital. Agreement between subject and collateral reports of drug and alcohol use in the past 30 days was high (85% to 95% agreement), although kappas were modest (.26 to .54). Notably, there were very few cases in which collaterals reported use that the subject had denied. Fewer than 2% of collaterals provided additional reports of drug use not obtained from self-report. On the basis of these results, the authors indicated that the time and expense necessary for obtaining collateral reports of drug use may be unwarranted in outpatient samples. However, this study assessed

lifetime, not current (past 12 months), substance use disorders, and drug and alcohol use in the past 30 days was relatively uncommon (e.g., 26% reported alcohol use). Therefore, the high rates of S-C agreement reported in this study might have been affected by the relatively infrequent use of substances in this sample, such that the generalizability of these findings to SMI samples with higher rates of current substance use remains to be established.

Drake et al. (1990) used multiple data sources, including collateral information from case managers, to reach consensus diagnoses regarding lifetime and current alcohol-related disorders in 75 psychiatric outpatients with schizophrenia-spectrum disorders. Although the results indicated that case manager and Diagnostic and Statistical Manual of Mental Disorders (3rd ed., rev.; American Psychiatric Association, 1987) interviewer ratings agreed on 88% of subjects' current alcohol diagnoses, the authors did not report on S-C agreement of recent drinking and drug use behavior. However, the results did indicate that case managers were able to make reliable and accurate assessments of current alcohol use diagnoses in persons with an SMI. Finally, Weiss et al. (2000) investigated the value of collateral reports of substance use for 32 psychiatric inpatients with current bipolar disorder and substance dependence. Collateral reports of substance use were compared to patient self-reports and urine toxicology screens. For the 132 assessments completed, there were 99 (75%) instances of agreement between collateral reports and the combined self-report and urine screen data (e.g., a positive collateral report and a positive self-report or urine screen). However, it is difficult to interpret this finding, because the percentage agreement statistic can be inflated by random or chance agreements (Hoffman & Ninonuevo, 1994).

The present study contributes significantly to the literature by investigating S-C agreement in a large sample of SMI individuals with current alcohol abuse or dependence. Although good consistency between subject and collateral reports has been shown across populations and settings, the vast majority of alcohol and drug treatment studies that have examined the accuracy of self-report have excluded individuals with bipolar or psychotic disorders. Consistent with recent research in the area of S-C reports (e.g., Stasiewicz et al., 1997), statistical techniques that take chance agreement into account were used. Moreover, this study investigated respondent characteristics (e.g., alcohol dependence, cognitive impairment, psychiatric symptoms) as well as collateral/relationship characteristics (e.g., collateral confidence, frequency of contact) predicted to affect the degree of consistency between subject and collateral reports among SMI individuals. Finally, this study examined agreement between a biochemical measure of drug use and reports of drug use from both subjects and collaterals. In particular, this study extends previous work (e.g., Fals-Stewart, O'Farrell, Freitas, McFarlin, & Rutigliano, 2000) by examining whether self- and collateral reports of drug use are consistent with urine drug test results for a sample of SMI individuals with a comorbid alcohol use disorder.

Method

Participants

The subjects were 207 men and women seeking outpatient dual-diagnosis treatment from a university-affiliated community mental health center. Forty (19%) subjects were excluded from analyses because of missing baseline collateral data.[1] The remaining sample (n = 167)

was 57% female and 43% male, with a mean age of 39.78 years (*SD* = 8.19) and a mean of 11.75 years of education (*SD* = 1.87). Most (68%) self-identified as African American; 24% self-identified as Caucasian, 4% as Latino, 2% as Native American, and 2% as of another ethnicity. The sample was mostly single (97%), unemployed (96%), and low income (84% reported annual incomes less than $10,000).

All potential subjects took part in an initial diagnostic screening session in which the Diagnostic Interview Schedule for *DSM–IV* (Robins, Cottler, Bucholz, & Compton, 1995) was administered. Subjects were eligible if they had lived at their current address for at least 6 months or could provide two persons as locators, scored at least 23 (with scores of 22 considered on a case-by-case basis) on the Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975) to ensure adequate cognitive functioning for study participation, and met Diagnostic and Statistical Manual of Mental Disorders (4th ed.; *DSM–IV;* American Psychiatric Association, 1994) criteria for a current (i.e., past 12 months) alcohol use disorder and a current schizophrenia-spectrum and/or bipolar disorder. Most (95%) met criteria for current alcohol dependence, while the remaining 5% met criteria for alcohol abuse. For current drug abuse or dependence, 76% met criteria for cocaine abuse or dependence, 53% met criteria for marijuana abuse or dependence, 20% met criteria for opiate abuse or dependence, 15% met criteria for sedative or hypnotic abuse or dependence, and 8% met criteria for amphetamine abuse or dependence. For comorbid mental disorder diagnoses, 21 (13%) subjects met criteria for a schizophrenia-spectrum disorder only (i.e., and not bipolar disorder), 95 (57%) subjects met criteria for bipolar disorder only (i.e., and not a schizophrenia-spectrum disorder), and 51 (30%) subjects met criteria for schizoaffective disorder (i.e., both schizophrenia-spectrum and bipolar disorders).

Procedure

Individuals were approached during the first 2 weeks following treatment entry to participate in a 6-month, naturalistic, longitudinal study examining predictors of relapse to alcohol and drug use. This study was approved by the university's institutional review board. Data for this report were derived from the diagnostic and baseline interviews conducted at the beginning of the subjects' treatment and the baseline telephone interviews with collaterals. Prior to every session, a breath test was administered to ensure that the subjects' blood alcohol level was zero. If a subject had a positive blood alcohol level, the session was rescheduled. Subjects were assured that their responses would be kept confidential, would not be discussed with treatment staff, and would not affect their treatment status. They received store gift cards ($15 at the diagnostic session and $30 at the baseline interview) for their participation. At the diagnostic session, subjects provided written informed consent, and their eligibility for the study was determined. At the baseline interview, held approximately 1 week later, subjects completed a variety of psychometrically sound questionnaires (described below). Since some SMI individuals have limited education and

[1]Baseline collateral data were missing for 40 subjects, either because the subjects refused to provide permission for collateral contact or because a collateral could not be reached. In supplementary analyses comparing subjects with (n = 167) and without (n = 40) collateral data, no demographic differences or differences in recent substance use were found. However, subjects with collateral data (M = 27.54, SD = 1.83) at baseline scored higher on the Mini-Mental State Exam than did participants without collateral data (M = 26.88, SD = 1.90), $t(204) = -2.05$, $p < .05$.

reading ability, trained research interviewers read all questionnaires and measures to subjects and recorded their responses.

Measures

At the diagnostic session, a comprehensive background questionnaire was used to obtain demographic characteristics, current status information (e.g., marital, employment, residential), and treatment history (e.g., number of psychiatric hospitalizations). In addition, the MMSE (Folstein et al., 1975) was administered at the diagnostic session. The MMSE is a widely used, brief cognitive screening measure that assesses orientation to time and place, attention, short-term memory, visual–spatial skills, and language functioning. The mean MMSE score was 27.54 (*SD* = 1.83; range = 22 to 30), with higher scores reflecting better cognitive functioning.

The Diagnostic Interview Schedule–IV (Robins et al., 1995) is a structured diagnostic interview used to obtain current and lifetime *DSM–IV* Axis I diagnoses. The following sections of the measure were administered by a trained research interviewer to all participants: Alcohol Use, Substance Use, Depression, Mania, and Schizophrenia. The Short Alcohol Dependence Data Questionnaire (SADD; Davidson & Raistrick, 1986) is a 15-item measure of alcohol dependence. The SADD has demonstrated good internal reliability (Raistrick, Dunbar, & Davidson, 1983) and concurrent validity (Davidson & Raistrick, 1986). In previous studies of SMI individuals (e.g., Gonzalez, Bradizza, Vincent, Stasiewicz, & Paas, 2007), the SADD has had good internal consistency. The mean SADD score was 21.32 (*SD* = 8.96), indicating high levels of dependence.

The Severity of Drug Dependence Questionnaire (SDS; Gossop et al., 1995) is a five-item measure of the psychological components of drug dependence. The SDS has demonstrated high internal consistency and good test–retest reliability (Gossop, Best, Marsden, & Strang, 1997; Gossop et al., 1995). In previous research with SMI individuals (e.g., Gonzalez et al., 2007), the SDS has had good internal consistency. The mean SDS score was 9.37 (*SD* = 4.38), indicating high levels of drug dependence.

The Structured Clinical Interview for the Positive and Negative Syndrome Scale (SCI-PANSS; Kay, Opler, & Fiszbein, 1992) is a 30-item measure of both positive (i.e., productive) symptoms and negative (i.e., deficit) symptoms of schizophrenia. Each item is scored on a 7-point severity scale (1 = *absent*, 7 = *extreme*), and items are summed for scoring. Of the 30 items, we administered 18 items, including the 7 items that reflect positive symptoms (e.g., hallucinations, delusions, excitement, hostility) and the 7 items that tap negative symptoms (e.g., blunted affect, poor rapport, emotional withdrawal, stereotyped thinking). On the basis of Shrout and Fleiss's (1979) ICC (3,1) formula for intraclass correlations (ICCs), interrater reliability for the Positive (ICC = .82) and Negative Symptoms (ICC = .71) subscales was fair to good. In a study of the psychometric properties of the SCI-PANSS among outpatients with schizophrenia and mood disorders, Purnine, Carey, Maisto, and Carey (2000) found support for the validity of the Positive and Negative Symptoms scales for this population.

The Timeline Follow-Back (TLFB; L. C. Sobell & Sobell, 1996) is a calendar-based retrospective recall interview of daily substance use. The TLFB was used to assess subjects' daily alcohol and drug use over the previous 60 days, as reported by both subjects and their collaterals. The TLFB has been found to be a reliable and valid method for use with psychiatric outpatients (Carey, 1997). It also has been administered to collaterals in several studies (e.g., Fals-Stewart et al., 2000; Stasiewicz et al., 1997; Stasiewicz & Stalker, 1999). For alcohol use, we examined total number of drinks (across 60 days), number of drinking days, number of heavy drinking days (defined as four or more standard drinks per day for women and five or more standard drinks per day for men), average number of drinks per drinking day, and average number of drinks per day. For drug use, we examined number of illicit drug (i.e., any illicit drug) use days, number of days the subject used marijuana, and number of days the subject used cocaine or crack.

A drug screen was performed on unsupervised urine samples collected from all subjects at baseline via the OnTrak TesTcup from Roche Diagnostics Corporation (Indianapolis, IN). This diagnostic test is intended for the simultaneous detection of drugs or drug metabolites in urine: in particular, amphetamines (1,000 ng/ml), cocaine metabolite (i.e., benzoylecgonine; 300 ng/ml), THC (i.e., marijuana; 50 ng/ml), and morphine (i.e., opiates; 300 ng/ml). Thirty-nine percent (n = 62) of the subjects tested positive for the presence of one or more of these drugs. Of these 62 subjects, 44 (71%) tested positive for cocaine, 28 (45%) tested positive for THC (marijuana), 7 (11%) tested positive for opiates, and 2 (3%) tested positive for amphetamines.

Collateral Interviews

Subjects provided names and telephone numbers for at least two collaterals and gave written permission for study personnel to contact them. Subjects and collaterals were assured that collateral information would be kept confidential. Trained study personnel conducted 10–15-min telephone interviews with collaterals within 2 weeks following subjects' baseline interviews. Use of telephone interviews to collect collateral data reflects current practice in the S-C literature, as face-to-face interviews with collaterals would be more costly and less practical. Collaterals were paid $10 for their participation.

On the basis of a collateral questionnaire developed for this study, collaterals were asked about their relationship to the subject, the frequency and nature (e.g., live together, work together) of their contact with the subject, and the subject's current substance use. To assess frequency of contact, collaterals were asked to report how frequently they had seen or talked to subjects during the past 2 months on a 6-point scale (1 = *daily*, 6 = *less than once a month*). Finally, collaterals rated their confidence in the accuracy of their reports on a 10-point scale (1 = *not at all confident*, 10 = *extremely confident*). A 60-day TLFB also was administered to collaterals to assess subjects' alcohol and drug use over the same time period as each subject's baseline TLFB.

Statistical Analyses

S-C agreement—Many studies on S-C agreement in the field of substance abuse report the percentage of S-C agreement. However, percentage agreement may be misleading,

because it can be inflated by chance agreement (cf. Hoffman & Ninonuevo, 1994). The kappa statistic is popular for examining agreement because it "corrects" for chance agreement. However, a disadvantage of kappa is that it declines as the base rate moves away from .50 and is lowest when the base rate nears .00 or 1.00 (Grove, Andreasen, McDonald-Scott, Keller, & Shapiro, 1981; Hoffman & Ninoneuvo, 1994). Yule's (1912) *Y* has been proposed as an alternative because it is relatively independent of the base rate for the variables being considered when the reports from two "fallible" raters are compared (Hoffman & Ninonuevo, 1994, p. 232; Spitznagel & Helzer, 1985; Yule, 1912). Thus, Yule's *Y* is preferable when one is examining agreement for items with a wide range of base rates. Like kappa, Yule's *Y* ranges from −1.00 to 1.00, with .50 or higher considered acceptable (Hoffman & Ninonuevo, 1994). For dichotomous variables, we report base rates, percentage agreement, kappa, and Yule's *Y*. Given the wide range of base rates among variables, agreement was based on Yule's *Y*. To determine S-C agreement for continuous variables, we used Shrout and Fleiss's (1979) ICC (1,1) formula to compute ICCs, which correct for chance agreement. Tests for the differences in the magnitude of the ICCs were conducted with Fisher's *z* tests. Guidelines for interpreting ICCs as an index of agreement were based on Cichetti (1994): below .40 = poor, .40–.59 = fair, .60–.74 = good, .75–1.00 = excellent.

Discrepancy analyses—To examine factors that influence S-C agreement for this population, we computed discrepancy scores by taking the absolute difference between subject and collateral reports for each of the following four variables: number of drinking days, number of heavy drinking days, total number of drinks, and number of illicit drug use days during the 60-day TLFB period. These variables were chosen to reflect both the quantity and frequency of alcohol use and the frequency of illicit drug use. Higher discrepancy scores reflected greater magnitudes of discrepancy (regardless of direction) between subject and collateral reports. Discrepancy scores were then simultaneously regressed onto predictor variables reflecting key subject and collateral/relationship characteristics chosen on the basis of prior research. We chose to enter predictor variables simultaneously because previous research provides little evidence regarding the nature of associations among variables that may influence S-C agreement. To reduce skew and normalize their distributions, we log-transformed discrepancy scores prior to analyses. Although variable transformations may make results harder to interpret, they are frequently recommended to improve the analysis (Tabachnick & Fidell, 2001).

Results

Collateral Relationship, Frequency of Contact, and Confidence Ratings

For the total sample, collaterals were 31% friends; 18% parents; 14% siblings; 13% current or former significant others; 10% other family members (e.g., aunt, in-law); 7% children of subjects; 2% spouses; 2% counselors, therapists, and case managers; and 3% others. Nearly three fourths (73%) reported that they had known the subject for at least 5 years. Average relationship length was 19.58 years (range = 1 month to 53 years; *SD* = 15.66 years). Over the past 2 months, 72% of collaterals reported seeing or talking to subjects daily or four to six times per week, with an additional 23% reporting at least weekly contact; nearly one

fourth (24%) of collaterals lived with subjects. Overall, the sample of collaterals was very confident in their reports about subjects' behavior (range = 4 to 10, *M* = 8.88, *SD* = 1.26).

Bipolar Only and Schizoaffective Groups

A first step was to examine S-C agreement by specific mental disorder diagnosis, focusing on subjects diagnosed with only bipolar disorder (*n* = 95) versus those diagnosed with schizoaffective disorder (*n* = 51). Because of the small subsample, subjects diagnosed with a schizophrenia-spectrum disorder but no bipolar disorder (*n* = 21) were excluded. For both continuous and dichotomous variables, S-C agreement did not differ between groups. In the absence of any differences between the bipolar only and schizoaffective groups, the remainder of the analyses focus on the total sample of 167 participants with collateral data at baseline.

S-C Agreement for Alcohol, Drugs, and Other Variables

Table 1 presents the agreement statistics for dichotomous variables using collateral questionnaire data. Results reveal acceptable S-C agreement for both marital and employment status as well as illicit drug use and cocaine use. As expected, the highest rates of agreement were found for marital and employment status. Variables that assessed alcohol use and marijuana use did not reach acceptable levels of agreement.

Table 2 presents ICCs and paired *t* test results for the continuous substance use variables based on TLFB data. Although participant and collateral reports were significantly correlated, the magnitude of the ICCs was low to moderate (*r*s = .18 to .40), reflecting poor to fair agreement. The two highest correlations were for number of days the subject used cocaine or crack (*r* = .40) and number of heavy drinking days (*r* = .37). Paired *t* test results revealed systematic differences: Subjects reported higher mean quantities of alcohol and greater frequency of drinking and drug use for the previous 60 days than did collaterals. Thus, in comparison to subjects' self-reports, collaterals consistently underestimated subjects' substance use.

The paired *t* test, a parametric statistic, requires the assumption of a normal distribution. Because of the moderate skew of the majority of TLFB variables, we also used a nonparametric method for comparison purposes, the Wilcoxon signed rank test for paired samples (Gibbons, 1993). Table 3 presents the results based on the Wilcoxon signed rank test, which examines whether the distribution of two variables is the same for two related samples, taking into account both direction and magnitude. A collateral underestimate was defined as an instance in which a collateral reported less substance use (e.g., fewer drinking days) than did the subject; a collateral overestimate was defined as an instance in which a collateral reported more substance use (e.g., more drinking days) than did a subject. Collateral same estimates reflected “ties,” the number of which might have been inflated by a preponderance of zero days of substance use for some of the substance use variables. For every variable examined, the two-tailed *Z* test was statistically significant (not shown in Table 3), indicating that the patterns of responses of subjects and their collaterals were significantly different. Consistent with the paired *t* test results, there was a higher percentage of collateral underestimates than of collateral overestimates. Thus, these results provide

important information regarding the direction of agreement between subjects and their collaterals.

Factors Affecting Agreement Between Subjects and Collaterals

To investigate which subject and collateral/relationship characteristics influenced the magnitude of agreement, we conducted four simultaneous multiple regression analyses. The dependent measure for each regression was the discrepancy score (i.e., absolute difference between subject and collateral reports) for number of drinking days, number of heavy drinking days, total number of drinks, and number of illicit drug use days.[2] Demographic predictors were sex (1 = male, 0 = female) and ethnicity (1 = European American, 0 = minority).[3] Other predictors were cognitive functioning (MMSE); positive and negative symptoms of schizophrenia (SCI-PANSS); alcohol dependence (SADD); drug dependence (SDS); urinalysis results (1 = positive, 0 = negative); and three collateral variables reflecting the nature of the relationship (1 = first-degree relative, 0 = not), frequency of contact, and collaterals' confidence in the accuracy of their reports.

Significant predictors of S-C discrepancies regarding substance use are summarized below. Sex and ethnicity were not related to three of the four discrepancy scores when we controlled for all other variables in the model. However, being European American (β=.22, $p < .05$) predicted greater discrepancy regarding number of illicit drug use days. Lower MMSE scores (reflecting poorer cognitive functioning) predicted greater discrepancy for total number of drinks ($\beta = -.25$, $p < .01$) and number of heavy drinking days ($\beta = -.17$, $p < .05$). Schizophrenic symptoms were unrelated to discrepancy scores. Higher SADD scores (reflecting greater alcohol dependence) predicted greater discrepancy for number of heavy drinking days ($\beta = .28$, $p < .001$). Collateral/relationship variables were not consistently related to S-C discrepancy when we controlled for all other variables. Less frequent S-C contact was related to greater discrepancy for total alcohol days ($\beta = .19$, $p < .05$). There was a nonsignificant trend for lower collateral confidence to be related to greater discrepancy for total number of drinks ($\beta = -.16$, $p = .07$). Further, a greater discrepancy for number of heavy drinking days was identified for collaterals who were not first-degree relatives ($\beta = -.20$, $p < .01$). The most consistent predictor of S-C discrepancies was urine drug test result. In particular, having a positive drug screen was significantly related to greater discrepancy for all four outcome variables (betas ranged from .25 to .42, $p < .01$). Overall, modest adjusted R^2 values (.18 to .20 for the four outcome variables) suggested that a large proportion of the variance in discrepancy outcomes remained to be explained.[4]

[2]Use of difference scores, including their use as dependent variables, has been a source of some debate. For a thorough discussion of the issues as well as a description of alternative approaches, please see Edwards (1995).

[3]Since age and education were unrelated to discrepancy scores in preliminary analyses, they were omitted as demographic predictors.

[4]The purpose of the absolute difference approach used here was to examine the relation between a variety of independent variables and the magnitude, but not the direction, of S-C agreement (i.e., nondirectional indexes). However, taking into account the direction of agreement may clarify the interpretation of the results. On the basis of an anonymous reviewer's suggestion, we reran the regression analyses without cases in which collaterals overestimated subjects' substance use (collateral > subject) to determine whether the pattern of findings remained the same for collateral underestimators (collateral < subject). Even when collateral overestimators were dropped from the analysis, a very similar pattern of results emerged. The only difference was that in the reanalysis, poorer cognitive functioning predicted greater S-C discrepancies for all four outcomes considered.

To illustrate the finding that being drug positive was associated with greater S-C discrepancy regarding substance use, we examined S-C agreement separately for the drug positive and drug negative groups. Table 4 presents S-C agreement for continuous variables comparing drug-negative and drug-positive groups. The correlations demonstrate that subjects who tested negative for drugs demonstrated better (i.e., “less worse”) agreement with collaterals than did subjects who tested positive. Among drug-negative subjects, ICCs were significant and low to moderate in magnitude (*r*s = .25 to .57), reflecting poor to fair agreement. In contrast, among drug-positive subjects, three correlations were significant (*r*s = .22 to .39) but reflected poor agreement, and the remaining nonsignificant correlations were low in magnitude (*r*s = .00 to .16), reflecting mostly poor agreement. It should be noted, though, that there was a significant difference between the ICCs for only one substance use variable. The ICC for average drinks per drinking day was significantly larger for the drug-negative group.

Agreement Between Subjects' Self-Reported Drug Use, Collaterals' Reports of Drug Use, and Urine Drug Test Results

Another goal of this study was to examine the convergent validity of subjects' and collaterals' day-by-day (i.e., TLFB) reports of drug use using a biochemical measure. A urine drug screen was performed at the start of subjects' baseline session, prior to questionnaire completion. Analyses of urine drug tests were limited to THC (marijuana) and cocaine (benzoylecgonine), the illicit drugs most frequently used by the sample. The urine detection time for cocaine (benzoylecgonine) is 2–4 days (Hawks & Chiang, 1986). On the basis of procedures used by Hersh, Mulgrew, Van Kirk, and Kranzler (1999), we examined urine detection windows of 3 and 5 days for cocaine. For instance, a 3-day window on the TLFB would include the day of the drug test (baseline session) as well as the 2 days prior. If cocaine use was reported on any of the 3 days, the TLFB report was judged positive for cocaine use.

The urine detection time for THC (cannabis) is 14–42 days (Hawks & Chiang, 1986). Fals-Stewart et al. (2000) systematically examined the large detection window for THC and determined that a 22-day window “provided the best balance between specificity and sensitivity” (p. 139); however, agreement between urine test results and self-report was not significantly improved by the use of a less than 42-day urine detection window versus a 42-day window (i.e., largest window). Therefore, to cover the full range of detection times, we used urine detection windows for THC of 14, 22, and 42 days. Thus, a 42-day window on the TLFB would include the day of the drug test (baseline session) as well as the 41 days prior.

We report agreement for self-reported drug use, collateral reports of drug use, and drug test results using percentage agreement, kappa, and Yule's *Y*. As shown in Table 5, subjects' TLFB reports of cocaine use exceeded acceptable levels of agreement (i.e., Yule's $Y \geq .50$) with drug test results for both the 3-day and the 5-day detection windows; however, collaterals' reports of cocaine use within the 5-day window just reached an acceptable level of agreement. A similar pattern was found for marijuana use (see Table 6): Subjects' TLFB reports exceeded acceptable levels of agreement with drug test results for every detection

window. However, collateral TLFB reports and drug test results just reached an acceptable level of agreement for the 14-day and 22-day detection windows and did not reach an acceptable level of agreement for the 42-day detection window. These findings confirm the accuracy of subjects' self-reported drug use on the TLFB but suggest that collaterals' TLFB reports of subjects' drug use had a lower degree of accuracy.

Results of 2 (urine drug test results) × 2 (participant TLFB report) × 2 (collateral TLFB report) crosstabs revealed some underreporting of cocaine and marijuana use by subjects on the TLFB.[5] Of the 44 subjects who tested positive for cocaine, 22 (50%) did not report cocaine use on the TLFB within the 3-day window, and 18 (41%) did not report cocaine use on the TLFB within the 5-day window. Of the 28 subjects who tested positive for marijuana, 11 (39%) did not report marijuana use on the TLFB within the 14-day window, 7 (25%) did not report marijuana use on the TLFB within the 22-day window, and 3 (11%) did not report marijuana use on the TLFB within the 42-day window. The crosstabs also revealed few, if any, cases (i.e., 0 to 2) in which the cocaine or marijuana urine test was positive, the subject did not report use of that drug on the TLFB within a given detection window, but the collateral reported subject use of the drug on the TLFB within the same detection window. Thus, collaterals' TLFB reports regarding drug use provided little additional information beyond that which the subject already had reported.

Discussion

This study is a comprehensive evaluation of factors associated with agreement between severely mentally ill alcohol abusers and collateral reports of alcohol and drug use, extending the previous S-C investigations of Carey and Simons (2000) and Weiss et al. (2000). Compared to previous studies investigating S-C agreement among non-SMI substance abusers (e.g., Babor et al., 2000), we found that S-C agreement was generally poor when we used recommended guidelines for interpreting ICCs (Cichetti, 1994) and other indexes (i.e., kappa and Yule's *Y*) of S-C agreement (Hoffman & Ninonuevo, 1994). In addition, S-C agreement appeared less favorable when the patient tested positive for drug use on the day of the assessment; ICCs were mostly nonsignificant and low in magnitude among drug-positive subjects and significant and low to moderate in magnitude among drug-negative subjects (although the differences between ICCs were nonsignificant for all but one variable). To assess the effect of respondent characteristics on S-C agreement, we regressed discrepancy scores for several substance use variables onto a set of predictors that included cognitive functioning, alcohol and drug dependence severity, symptoms of schizophrenia, and several collateral/relationship variables. Overall, collateral variables showed no consistent relation to discrepancy scores. However, less frequent contact was related to greater disagreement regarding the frequency of alcohol use. Similarly, Stasiewicz et al. (1997) found that frequency of contact was related to S-C agreement regarding alcohol

[5]We also examined the extent to which subjects reported drug use on the TLFB but had negative drug tests (i.e., overreporting). For cocaine, 4 subjects reported use on the TLFB within the 3-day window but had a negative drug test. Two subjects reported cocaine use on the TLFB within the 5-day window but had a negative drug test. For marijuana, 2 subjects reported use within the 14-day window but had a negative drug test for THC, 6 subjects reported use within the 22-day window but had a negative drug test, and 19 reported use within the 42-day window but had a negative drug test. Thus, subjects tended not to overreport their cocaine and marijuana use on the TLFB.

use among dually diagnosed individuals. However, using another dually diagnosed sample, Carey and Simons (2000) found that frequency of contact was related to agreement regarding drug use but not alcohol use. There was a nonsignificant trend for higher collateral confidence to be related to greater S-C agreement for one of the four discrepancy outcomes. The predictive utility of collateral confidence might have been hampered by (a) use of a single-item (i.e., less reliable) measure and (b) the restricted variability observed for this measure. Also, several studies (e.g., L. C. Sobell, Agrawal, & Sobell, 1997) have demonstrated that S-C agreement was enhanced when spouse or partner collaterals were used. However, for certain subject populations, such as the SMI individuals studied here, spouse or partner collaterals are not available (Carey & Simons, 2000). In these populations, persons with close relationships and frequent contact with the subject, such as first-degree relatives, may prove to be a superior source of collaterals. Indeed, being a first-degree relative was related to greater S-C agreement for one of the discrepancy variables.

After we controlled for other respondent characteristics as well as collateral/relationship variables, urine drug screen results were predictive of all four outcomes considered. Thus, consistent with previous literature, state of sobriety was detrimental to the validity of self-reports of substance use in SMI individuals with a comorbid alcohol use disorder. Indeed, both a positive drug test and higher levels of substance use have been shown in previous studies to be related to poorer convergent validity (e.g., Stasiewicz et al., 1997; M. B. Sobell, Sobell, & Vanderspek, 1979).

After we controlled for other respondent characteristics and collateral/relationship variables, schizophrenic symptoms were not associated with S-C agreement. However, worse cognitive functioning was related to lower S-C agreement for two of the four outcomes considered (and when we reran the analysis without collateral overestimators, worse cognitive functioning was related to lower S-C agreement for all four outcomes considered). We found four previous studies that investigated the relationship between cognitive functioning and S-C agreement. In two of the studies, no relationship was found between cognitive functioning and discrepancy scores (Stasiewicz et al., 1997; Stasiewicz & Stalker, 1999). The subjects in those studies were not severely mentally ill, and the MMSE was not used to evaluate cognitive functioning. In the third study, of 13 alcohol abusers in treatment (Miller & Barasch, 1985), greater cognitive impairment (as assessed by the MMSE) was associated with greater disagreement between subject and collateral reports of alcohol consumption. Finally, in a large sample (N = 1,726) of men and women with alcohol abuse or dependence, Babor et al. (2000) found greater S-C disagreement in clients with greater cognitive impairment. Therefore, whereas cognitive impairment has been hypothesized to affect S-C agreement, results are mixed in studies examining substance abusers without SMI. However, the results of the present study indicate that cognitive functioning is relevant to S-C agreement in substance abusers with SMI.

Similar to previous studies with SMI individuals, we found that collaterals rarely provided more information than was provided by the subjects themselves, which calls into question the value of routine use of collateral informants. As discussed by others (Carey & Simons, 2000; Connors & Maisto, 2003), the tendency for subjects to provide more information only has been observed in studies in which the subject knows that the collateral is being

contacted. The extent to which the overreporting relative to collaterals reported in the literature will generalize to situations in which collaterals are not contacted is unknown. As Carey and Simons (2000) pointed out, the value of collateral information may reside in a metacommunication to the subject that self-reports can be corroborated and, therefore, inaccurate self-reports would be caught. Research is needed to explore this possibility.

This pattern of results also was confirmed with the urine drug screen results as the corroborating source. We found high agreement between subjects' self-reported substance use and urine drug screen results taken at the baseline interview. For both marijuana and cocaine, Yule's *Y* ranged from .68 to .82 across the varied urine detection times for these drugs. These results are consistent with comparable findings with urinalysis as the corroborating source and percentage agreement as the statistic used to assess S-C correspondence (Weiss et al., 1998, 2000). Agreement was lower but still acceptable when we compared subjects' drug test results to collaterals' TLFB reports.

Although the focus of this study is on factors that predict the extent of S-C agreement, there were some differences in agreement across the various dependent variables. Notably, number of heavy drinking days showed a relatively high ICC, and this was the only dependent variable for which the paired *t* test between subject and collateral was not statistically significant, which suggests relatively higher S-C agreement for this variable. Further, this variable showed the largest number of significant associations with the predictor variables. In particular, greater S-C discrepancy for number of heavy drinking days was significantly related to greater cognitive impairment, greater severity of alcohol dependence, and having a collateral who was not a first-degree relative. Thus, for number of heavy drinking days, we have some knowledge of the conditions under which agreement for this variable may be enhanced.

Limitations of this study include that the urine drug test results should be considered preliminary because they were not conducted via immunoassay techniques, which may provide greater accuracy. Positive drug test results do not indicate the degree of intoxication but rather assess the presence or absence of a drug or drug metabolites in urine. Furthermore, urine drug testing may not detect low-level use (Hersh et al., 1999), and urine detection windows vary among individuals, which suggests that choice of a different window can result in different findings for agreement (Ehrman & Robbins, 1994). The relatively high level of agreement between subject self-report and urine drug test result might have been due to subjects' knowledge that urine drug testing would be conducted (Ehrman & Robbins, 1994). Also, our analyses examining agreement for continuous substance use variables, such as number of drinking days, were not sensitive to temporal patterns of consumption (Tucker, Vuchinich, Harris, Gavornik, & Rudd, 1991). In other words, agreement statistics did not reflect when during the 60-day timeline subjects and collaterals reported use. Although such analyses are beyond the scope of this article, future research might examine whether agreement varies on the basis of when use occurred. Finally, we had modest success in predicting agreement using key variables identified in the existing literature. Clearly, further research is needed to identify other factors that may help explain variability in agreement. Although the present study incorporated procedures predicted to enhance S-C agreement (e.g., confidentiality assurances, subject knowledge of

collateral contact, clear instructional set), agreement was still modest at best. One untapped source of potential variability in S-C agreement is the collateral respondent. For example, collaterals (e.g., friends) nominated by dually diagnosed individuals may be substance users themselves. Therefore, collaterals' substance use and attitudes toward drinking may be useful to assess. As Connors and Maisto (2003) pointed out, “as knowledge about characteristics of both subjects and collaterals that predict consistency of their reports advances, clinical researchers will be better judges of the accuracy of self-reports of alcohol use” (p. 28).

In conclusion, the present study demonstrates that S-C agreement in reports of drinking and drug use in individuals with comorbid alcohol abuse or dependence and an SMI was somewhat less favorable compared to the results of previous studies investigating S-C agreement in other clinical populations. In addition, collateral reports rarely provided more information than was provided by the subjects themselves. As others have suggested, future research might investigate interview procedures designed to increase the validity of self-report information (e.g., Babor et al., 2000; Laforge, Borsari, & Baer, 2005).

Somewhat more encouraging, the present study found that the degree of S-C agreement regarding substance use consistently improved when individuals who screened positive for one or more illicit drugs were removed from the analyses. Given the high rates of other substance use disorders (i.e., cocaine and marijuana abuse and dependence) in this sample and the less favorable S-C agreement of drug-positive subjects, urine drug screens should be administered prior to collection of reports of substance use by SMI individuals. As the results of this study show, agreement is highest when subjects' self-reports of drug use are corroborated by urine drug test results.

Acknowledgments

This research was funded by National Institute on Alcohol Abuse and Alcoholism Grant R01 AA12805 to Clara M. Bradizza. We thank Gregory G. Homish and Paula Richards for their help with data analysis and Adam B. Stasiewicz for his helpful comments on a draft of this article.

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3rd ed.. Author; Washington, DC: 1987. rev.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed.. Author; Washington, DC: 1994.

Babor TF, Brown J, Del Boca FK. Validity of self-reports in applied research on addictive behaviors: Fact or fiction? Behavioral Assessment. 1990; 12:5–31.

Babor, TF.; Del Boca, FK. Just the facts: Enhancing measurement of alcohol consumption using self-report methods. In: Litten, RZ.; Allen, JP., editors. Measuring alcohol consumption: Psychosocial and biochemical methods. Humana Press; Totowa, NJ: 1992. p. 3-19.

Babor TF, Steinberg K, Anton R, Del Boca F. Talk is cheap: Measuring drinking outcomes in clinical trials. Journal of Studies on Alcohol. 2000; 61:55–63. [PubMed: 10627097]

Babor TF, Stephens RS, Marlatt GA. Verbal report methods in clinical research on alcoholism: Response bias and its minimization. Journal of Studies on Alcohol. 1987; 48:410–424. [PubMed: 3312821]

Carey KB. Reliability and validity of the time-line follow-back interview among psychiatric outpatients: A preliminary report. Psychology of Addictive Behaviors. 1997; 11:26–33.

Carey KB, Correia C. Severe mental illness and addictions: Assessment considerations. Addictive Behaviors. 1998; 23:735–748. [PubMed: 9801713]

Carey KB, Simons J. Utility of collateral information in assessing substance use among psychiatric outpatients. Journal of Substance Abuse. 2000; 11:139–147. [PubMed: 10989774]

Cichetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment. 1994; 6:284–290.

Connors GJ, Maisto SA. Drinking reports from collateral individuals. Addiction. 2003; 98(Suppl. 2): 21–29. [PubMed: 14984239]

Davidson R, Raistrick D. The validity of the Short Alcohol Dependence Data (SADD) Questionnaire: A short self-report questionnaire for the assessment of alcohol dependence. British Journal of Addiction. 1986; 81:217–222. [PubMed: 3458489]

Drake RE, Osher FC, Noordsy DL, Hurlbut SC, Teague GB, Beaudett MS. Diagnosis of alcohol use disorders in schizophrenia. Schizophrenia Bulletin. 1990; 16:57–67. [PubMed: 2333482]

Drake RE, Osher FC, Wallach M. Alcohol use and abuse in schizophrenia: A prospective community study. Journal of Nervous and Mental Disease. 1989; 177:408–414. [PubMed: 2746194]

Edwards JR. Alternatives to difference scores as dependent variables in the study of congruence in organizational research. Organizational Behavior and Human Decision Processes. 1995; 64:307–324.

Ehrman RN, Robbins SJ. Reliability and validity of 6-month timeline reports of cocaine and heroin use in a methadone population. Journal of Consulting and Clinical Psychology. 1994; 62:843–850. [PubMed: 7962889]

Fals-Stewart W, O'Farrell TJ, Freitas TT, McFarlin SK, Rutigliano P. The timeline follow-back reports of psychoactive substance use by drug-abusing patients: Psychometric properties. Journal of Consulting and Clinical Psychology. 2000; 68:134–144. [PubMed: 10710848]

Folstein MF, Folstein SE, McHugh PR. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research. 1975; 12:189–198. [PubMed: 1202204]

Gibbons, JD. Nonparametric statistics: An introduction. Sage; Newbury Park, CA: 1993.

Gonzalez VM, Bradizza CM, Vincent PC, Stasiewicz PR, Paas ND. Do individuals with a severe mental illness experience greater alcohol and drug-related problems? A test of the supersensitivity hypothesis. Addictive Behaviors. 2007; 32:477–490. [PubMed: 16828977]

Gossop M, Best D, Marsden J, Strang J. Test-retest reliability of the Severity of Dependence Scale. Addiction. 1997; 92:353. [PubMed: 9219397]

Gossop M, Darke S, Griffiths P, Hando J, Powis B, Hall W, et al. The Severity of Dependence Scale (SDS): Psychometric properties of the SDS in English and Australian samples of heroin, cocaine and amphetamine users. Addiction. 1995; 90:607–614. [PubMed: 7795497]

Grove WM, Andreasen NC, McDonald-Scott P, Keller MB, Shapiro RW. Reliability studies of psychiatric diagnosis: Theory and practice. Archives of General Psychiatry. 1981; 38:408–413. [PubMed: 7212971]

Hawks, RL.; Chiang, CN. Examples of specific drug assays. In: Hawks, RL.; Chiang, CN., editors. Urine testing for drugs of abuse. National Institute on Drug Abuse; Rockville, MD: 1986. p. 84-112.Research Monograph No. 73

Hersh D, Mulgrew CL, Van Kirk J, Kranzler HR. The validity of self-reported cocaine use in two groups of cocaine abusers. Journal of Consulting and Clinical Psychology. 1999; 67:37–42. [PubMed: 10028207]

Hoffman NG, Ninonuevo FG. Concurrent validation of substance abusers' self-reports against collateral information: Percentage agreement vs. κ vs. Yule's. Y. Alcoholism: Clinical and Experimental Research. 1994; 18:231–237.

Kay, SR.; Opler, LA.; Fiszbein, A. Positive and Negative Syndrome Scale: Manual. Multi-Health Systems, Inc.; North Tonawanda, NY: 1992.

Laforge RG, Borsari B, Baer JS. The utility of collateral informant assessment in college alcohol research: Results from a longitudinal prevention trial. Journal of Studies on Alcohol. 2005; 66:479–487. [PubMed: 16240555]

Miller F, Barasch A. The under-reporting of alcohol use: The role of organic mental syndromes. Drug and Alcohol Dependence. 1985; 15:347–351. [PubMed: 4053972]

Purnine DM, Carey KB, Maisto SA, Carey MP. Assessing positive and negative symptoms in outpatients with schizophrenia and mood disorders. Journal of Nervous and Mental Disease. 2000; 188:653–661. [PubMed: 11048814]

Raistrick D, Dunbar G, Davidson R. Development of a questionnaire to measure alcohol dependence. British Journal of Addiction. 1983; 78:89–95. [PubMed: 6573181]

Robins, LN.; Cottler, LB.; Bucholz, KK.; Compton, WM. Diagnostic Interview Schedule for DSM–IV (DIS–IV). National Institute of Mental Health; St. Louis, MO: 1995.

Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin. 1979; 86:420–428. [PubMed: 18839484]

Sobell LC, Agrawal S, Sobell MB. Factors affecting agreement between alcohol abusers' and their collaterals' reports. Journal of Studies on Alcohol. 1997; 58:405–413. [PubMed: 9203122]

Sobell, LC.; Sobell, MB. Timeline Follow-Back user's guide. Addiction Research Foundation; Toronto, Ontario, Canada: 1996.

Sobell MB, Sobell LC, Vanderspek R. Relationships among clinical judgment, self-report, and breath-analysis measures of intoxication in alcoholics. Journal of Consulting and Clinical Psychology. 1979; 47:204–206. [PubMed: 429662]

Spitznagel EL, Helzer JE. A proposed solution to the base-rate problem in the kappa statistic. Archives of General Psychiatry. 1985; 42:725–728. [PubMed: 4015315]

Stasiewicz PR, Bradizza CM, Connors GJ. Subject-collateral reports of drinking in inpatient alcoholics with comorbid mental disorders. Alcoholism: Clinical and Experimental Research. 1997; 21:530–536.

Stasiewicz PR, Stalker RG. Subject-collateral reports of drinking in inpatient alcoholics with comorbid cocaine dependence. American Journal of Drug and Alcohol Abuse. 1999; 25:319–329. [PubMed: 10395163]

Tabachnick, BG.; Fidell, LS. Using multivariate statistics. 4th ed.. Allyn & Bacon; Boston: 2001.

Toneatto T, Sobell LC, Sobell MB. Predictors of alcohol abusers' inconsistent self-reports of their drinking and life events. Alcoholism: Clinical and Experimental Research. 1992; 16:542–546.

Tucker JA, Vuchinich RE, Harris CV, Gavornik MG, Rudd EJ. Agreement between subject and collateral verbal reports of alcohol consumption in older adults. Journal of Studies on Alcohol. 1991; 52:148–155. [PubMed: 2016875]

Weiss RD, Greenfield SF, Griffin ML, Najavits LM, Fucito LM. The use of collateral reports for patients with bipolar and substance use disorders. American Journal of Drug and Alcohol Abuse. 2000; 26:369–378. [PubMed: 10976663]

Weiss RD, Najavits LM, Greenfield SF, Soto JA, Shaw SR, Wyner D. Validity of substance use self-reports in dually diagnosed outpatients. American Journal of Psychiatry. 1998; 155:127–128. [PubMed: 9433351]

Yule GU. On the methods of measuring association between two attributes. Journal of the Royal Statistical Society. 1912; 75:581–642.

Table 1

Subject–Collateral Agreement for Dichotomous Variables

Variable	Base rate	% agreement	κ	Yule's *Y*
Marital status	.19	.90	.66	**.72**
Employment status	.04	.96	.48	**.76**
Alcohol use	.75	.60	.25	.40
Any illicit drug use	.68	.57	.25	**.53**
Marijuana use	.30	.73	.22	.38
Cocaine/crack use	.54	.66	.34	**.60**

Note. Because of missing data on some variables, sample sizes range from 116 to 167. Subject substance use variables are from the Timeline Follow-Back data. Collateral variables are from the collateral questionnaire. Boldface values of Yule's *Y* (≥ .50) reflect acceptable subject-collateral agreement.

Table 2

Subject–Collateral Agreement for Continuous Variables

		Subject		Collateral			
Substance use variable	**ICC**	***M***	***SD***	***M***	***SD***	**Paired *t* value**	***df***
Total no. drinks	.31***	93.58	130.53	45.49	89.77	4.65***	156
No. of drinking days	.35***	13.55	17.16	8.56	16.38	3.45***	166
No. of heavy drinking days	.37***	8.99	12.87	6.69	15.29	1.76	163
Ave. drinks per drinking day	.18	9.39	6.84	6.57	4.53	2.83**	56
Ave. drinks per day	.25***	1.67	2.36	0.95	1.97	3.85***	160
No. of days used any illicit drugs	.23**	11.43	16.52	3.74	8.43	6.05***	161
No. of days used marijuana	.31***	1.72	4.16	0.80	2.95	2.76**	153
No. of days used cocaine/crack	.40***	5.10	8.44	2.80	6.762	3.89***	156

Note. Because of missing collateral data, the sample size for average drinks per drinking day is 57. For the remaining variables, sample sizes range from 154 to 167. All variables are from the subject and collateral Timeline Follow-Back data. ICC = intraclass correlation for subject–collateral reports; Ave. = average.

** $p < .01$.

*** $p < .001$.

Table 3

Percentages of Collateral Underestimates, Same Estimates, and Overestimates Based on Wilcoxon Signed Rank Tests

Substance use variable	Collateral underestimate	Collateral same estimate	Collateral overestimate	*n*
Total no. of drinks	61	22	17	157
No. of drinking days	59	24	17	167
No. of heavy drinking days	49	35	16	164
Ave. drinks per drinking day	68	0	32	57
Ave. drinks per day	61	21	18	161
No. of days used any illicit drugs	58	32	10	162
No. of days used marijuana	25	69	6	154
No. of days used cocaine/crack	44	46	10	157

Note. Collateral underestimate = collateral < subject; collateral same estimate = collateral–subject tie; collateral overestimate = collateral > subject. Data in the last column represent the number of subject–collateral pairs. Ave. = average.

Table 4

Subject–Collateral Agreement by Urine Drug Screen Result (Continuous Variables)

	Positive drug screen		Negative drug screen	
Substance use variable	**ICC**	***n***	**ICC**	***n***
Total no. of drinks	.22*	56	.29**	94
No. of drinking days	.16	62	.39***	97
No. of heavy drinking days	.23*	59	.45***	97
Ave. drinks per drinking day[a]	.00	30	.57***	25
Ave. drinks per day	.15	58	.25**	95
No. of days used illicit drugs	.03	58	.32***	96
No. of days used marijuana	.16	50	.38***	96
No. of days used cocaine/crack	.39***	56	.38***	94

Note. Drug test results are based on subjects' drug urine screens at baseline. All variables are from collateral and subject Timeline Follow-Back data. ICC = intraclass correlation for subject–collateral reports; Ave. = average.

[a]Variable for which the correlations are significantly different at $p < .05$.

*$p < .05$.

**$p < .01$.

***$p < .001$.

Table 5

Agreement Between Urine Drug Test Result and Timeline Follow-Back Data for Cocaine Use

	3-day detection window			5-day detection window		
Reporter of cocaine use	**% agreement**	**κ**	**Yule's *Y***	**% agreement**	**κ**	**Yule's *Y***
Subject	.84	.53	**.68**	.87	.65	**.80**
Collateral[a]	.72			.77	.30	**.53**

Note. n = 159 subjects with urine drug test results. Urine drug test results were coded as positive/negative for cocaine use. Timeline Follow-Back data were coded yes/no for cocaine use within each detection window. Boldface values of Yule's *Y* (≥ .50) reflect acceptable subject-collateral agreement.

[a]Values for kappa and Yule's *Y* could not be computed for the 3-day detection window because there were no collateral reports of subject cocaine use within the 3-day window (i.e., cell counts = 0).

Table 6

Agreement Between Urine Drug Test Result and Timeline Follow-Back Data for Marijuana Use

	14-day detection window			22-day detection window			42-day detection window		
Reporter	**% agreement**	**κ**	**Yule's *Y***	**% agreement**	**κ**	**Yule's *Y***	**% agreement**	**κ**	**Yule's *Y***
Subject	.92	.68	**.82**	.92	.71	**.78**	.86	.61	**.75**
Collateral	.84	.29	**.53**	.84	.31	**.52**	.83	.31	.46

Note. n = 159 subjects with urine drug test results. Urine drug test results were coded positive/negative for marijuana use. Timeline Follow-Back data were coded yes/no for marijuana use within each detection window. Boldface values of Yule's Y (≥ .50) reflect acceptable subject–collateral agreement.