

Self-Report Stability of Adolescent Cigarette Use
Across 10 Years of Panel Study Data

Audrey M. Shillington*

Mark B. Reed

John D. Clapp

San Diego State University, School of Social Work, College of Health and Human
Services

Center for Alcohol and Other Drug Studies and Services

Address correspondence to AM Shillington, Ph.D., Center for Alcohol and
Other Drug Studies and Services, 6386 Alvarado Court, Ste. 224., San Diego
CA 92120. This study was funded by a grant from the National Institute on
Alcohol Abuse and Alcoholism 5R21AA016769, AM Shillington, P.I. 619-
229-2336 (phone) email: ashillin@mail.sdsu.edu

Abstract

This study is the first to examine adolescent cigarette report stability over ten years. Six waves of data were utilized from the National Longitudinal Survey of Youth. This study examined internal/logical consistency and external consistency. Report stability was higher for lifetime use reports than the age of onset reports. Wave-by-wave differences revealed stability increased across time with one third denying use in the first two wave comparisons but dropping to twenty percent by the last comparison. Overall, report agreement was higher for females, older adolescents, and Non-Hispanic/Non-Black youth. Implications regarding misclassification of users for prevention programs and measurement issues are discussed.

Key Words: Measurement, Report Stability, Agreement rates, reliability, Cigarettes

Background

Most of the data on national trends, prevalence and incidence of cigarette use among adolescents rely on self-reports of use. The importance of self-report data is considerable. There is concern about the reliability and validity of self-reported information with regard to all behaviors and especially so for illegal or socially unacceptable behaviors. Although cigarette smoking is not an illegal activity for adults, it is illegal for minors—that is, for children and adolescents (O'Malley, Bachman, & Johnston, 1983; Bailey, Flewelling, & Rachal, 1992). When being queried regarding behaviors that are illegal or socially stigmatizing there is the potential of under-reporting.

As Rice, Rochberg, Endicott, Lavori, & Miller (1992) discuss, a related paradigm to test-retest reliability is a test of temporal stability, which requires two interviews spaced far apart in time to determine how well respondents repeat their answers from one interview to another. Unlike the short temporal interval employed in test-retest reliability analyses, measures in stability studies are typically conducted one year or more after the initial survey period. It is assumed that agreement from Time 1 to Time 2 is indicative of a true lifetime clinical state. Reliability on the other hand is used to investigate the repeatability of responses to an instrument (Rice, et al., 1992).

A change in reporting lifetime cigarette use is considered the stability of a response over longer periods of time compared to more traditional test-retest reliability. It is conceptually different than the concept of reliability as well—in that reliability relates to the psychometric properties of a measure. The

underlying assumption of the concept of stability is that if one reports use at time 1, then lifetime use should be reported by the same respondent at every subsequent measurement period. Where test-retest reliability is aimed at measuring the potential error in a measure (Pedhazur & Schmelkin, 1991), stability is focused more on participant recall errors.

There have been a few studies that have examined the reliability, validity and/or stability of self-reports of alcohol and other drug use (AOD) use among adolescents using general population samples (O'Malley et al., 1983; Bailey et al., 1992; Barnea, Rahav, & Teichman, 1987; Smith, McCarthy, & Goldman, 1995; Johnson & Mott, 2001; Stanton, Papandonatos, Lloyd-Richardson & Niaura, 2007). These studies have yielded mixed findings with various methodological limitations.

There are two key pieces of information gathered through self-report surveys in epidemiological studies that can be effectively subjected to stability analyses. Such data would include variables that are reported as lifetime behaviors or the age of onset variables of particular behaviors. So, the first is "ever use" which asks a respondent if he or she used a particular substance ever in his or her lifetime. The second is "age of onset" which is asked if the individual reports use of a particular substance and asks respondents to report the specific age at which the substance was first used. These types of behaviors are different from those that one would expect to vary over time such as quantity or frequency reports or variables that query past three month use patterns.

Ever use

Ever use has been more widely studied in this literature compared to studies examining age of onset measures. Fendrich and Rosenbaum (2003) reported recanting rates (reporting use at time 1 but denying use at time 2) of 45% for alcohol use and about 50% for cigarette use among teens in a longitudinal study examining the effectiveness of the DARE program. Siddiqui, Mott, Anderson & Flay (1999) found a 24% report inconsistency rate for at least one substance. Johnson, Gerstein, & Rasinski, (1997) found a 25% decrease in reported alcohol incidence and an 18% decrease in reported marijuana incidence over a five year time span. Fendrich & Kim (2001) examined adult characteristics over a longitudinal design in 1988, 1992, and 1994. Specifically, they found 40% of cocaine users and 30% of marijuana users denied use in at least one follow up period after prior reports of use.

One possible explanation for the differences found between the studies showing a lack of report reliability and those demonstrating higher reliabilities may be attributed to how the data are coded. Bailey et al. (1992) found that by adjusting for consistent non-users (those reporting no use at both time 1 and time 2), rates of agreement between measurement periods dropped significantly. In other words, if the subjects who consistently report “no use” at time one and time two are retained in the calculations, then the percentage of agreement is much higher. However, if the adjustment such as that recommended by Bailey et al., (1992) is utilized in which the consistent non-users are excluded then agreement rates drop. For instance, on measures of lifetime use for marijuana, the agreement rate between time 1 and time 2 decreased from 95.6% to 83.2%

(Bailey, et al., 1992). Thus, for the study of report stability or consistency over time the adjusted percentage is a more accurate estimate of those who change their reports of lifetime use or age of onset. Changes from non-use to use from one wave to another are reported as incidence cases.

Respondent characteristics have been examined in prior research in order to identify variables associated with recanting or denial of previously reported behavior. Ethnicity is one of the more consistently found variables to be related to recanting. For example, several authors (Fendrich and Kim, 2001; Johnson & Mott, 2001; Shillington & Clapp, 2000; Stanton et al., 2007) have found higher rates of inconsistent reporting among African Americans and Hispanics as compared to Non-Hispanic, Non-African Americans. In fact Stanton and colleagues (2007) found that recanting was higher among ethnic minority groups even after controlling for other socio-demographic variables. Additionally, Shillington & Clapp (2000) found females were more consistent than males in cigarette ever-use reports across two waves of data collection. Such differences have been replicated by Percy, McAlister, Higgins, McCrystal & Thornton (2005) among young adolescents. However, these respondent characteristics have not been well studied among adolescent cigarette users since most have used adult samples.

Age at Onset

The second area of research examined in this literature concerns age of onset of use for particular substances. Understanding the reliability and the stability of reports of age of first use has not been firmly established (Johnson &

Mott, 2001) but is important for several reasons. First, age of onset may be more prone to measurement error compared to “ever use” questions because questions of this type are asking for more precise information compared to asking if one has ever engaged in a particular behavior. Second, age of onset is often utilized by researchers and policy makers alike as a marker variable for those who may be at more risk for future substance use (i.e., earlier onset is typically associated with heavier use).

One concept related to changes in age of onset is referred to as telescoping. Forward telescoping is the tendency to increase one’s age of onset as one ages chronologically. Backward telescoping is the opposite--the tendency to report one’s age of onset as younger as one gets older. The existence of telescoping in reported age of onset for substance use has direct implications for prevention providers, epidemiologists and policy analysts. Theories of substance use trajectories such as the Gateway Theory posit that individuals who begin to use alcohol and tobacco at early ages are at greater risk to progress to the use of illicit substances as well as an increased risk of progressing to problem use (Funkhouser, Goplerud, & Bass, 1992; Grant, 1998; Grant & Dawson, 1997; Hawkins, Catalano, & Miller, 1992; Kandel, Yamaguchi, & Chen, 1992; Mills & Noyes, 1984). Furthermore, the unreliability of age at onset for substance use limits the ability to: 1) predict individual outcomes, 2) accurately examine longitudinal trajectories in substance use patterns and problems, and 3) determine the appropriate timing to introduce prevention programs (Golub, Johnson, & Labouvie, 2000a).

To date, few studies have examined the psychometric properties of the important variable of age of onset. The reliability of age at onset for alcohol and drug use is mixed. Some studies show the reliability of age of onset as being average to good (Andreasen, et al., 1981; Fendrich, Weissman, Warner, & Mufson, 1990) and another study found the age of onset for drug use to have acceptable reliability (Prusoff, Merikangas, & Weissman, 1988). Bailey and associates (1992) however, found that measures concerning age of first use were not stable for any substance with agreements between waves near or below 35%. They note that age of onset report consistency is particularly high for substances with lower prevalence rates. Prause, Dooley, Ham-Rowbottom & Emptage (2007) reported that among adolescents, age of onset for alcohol use to has low reliability. Forward telescoping was found in several studies for substance use (Fendrich, et al., 1990; Golub, et al., 200a; Johnson & Schultz, 2005) and psychiatric disorders (Prusoff et al., 1988).

Internal Consistency

When cross-sectional data are used to assess the reliability of self-reports, authors have used the terms internal consistency, internal validity and logical consistency as a proxy for reliability (O'Malley et al., 1983; Barnea et al., 1987). Each of these terms relates to the consistency of respondents' answers to the same item or logically similar items within a cross-sectional interview or self-administered questionnaire (Barnea et al., 1987). Using this approach, a few studies have reported high levels of agreement between self-reports on logically

similar questions concerning the use of several types of substances (Barnea, et al., 1987; Bailey et al., 1992; Shillington & Clapp, 2000).

Limitations of Current Research

Nearly all studies examining the reliability of reports or longer-term stability of reports have been comparisons across two waves of data spaced 1-3 years apart (Fendrich et al., 1995; Fendrich & Vaughn, 1994; Golub et al., 2000b; Bailey, 1992; Smith, McCarthy, & Goldman, 1995; O'Malley, 1993; Shillington & Clapp, 2000). Longer time spans over four or more waves of data have been examined by only a handful of studies with time periods ranging up to 10 years (Johnson & Mott, 2001; Fendrich & Rosenbaum, 2003; Fendrich & Kim, 2001). Shillington and colleagues (1995) examined the stability of reports across ten years but with only two waves of data. There is a paucity of studies that examine report stability over longer periods of time and across more than two waves of data. The few that have used multiple waves have focused primarily on report stability of use rather than age of onset.

Cigarette Use among Youth

Rates of smoking prevalence among youth have decreased substantially since the late 1990s as a result of the Tobacco Settlement Agreement and subsequent banning of cigarette advertising targeting children as well as higher cigarette prices and excise taxes (Rosen & Maurer, 2008). Since 2003 these steep declines have leveled off and lifetime smoking prevalence rates among high school seniors was at 46.2% as of 2007 and 21.6% for past month use (Johnston, O'Malley, Bachman, & Schulenberg, 2008). Despite the large

decreases in adolescent smoking prevalence observed during the late 1990s 3 million adolescents between the ages of 12 and 17 smoke cigarettes (CDC, 2005) and half of those adolescents addicted to nicotine will eventually die from smoking-related diseases (Peto et al., 2004).

The majority (nearly 90%) of adult cigarette smokers initiated smoking during their preadolescent or adolescent years (CDC, 2006; NIDA 1997; Mosbach & Leventhal, 1988). In fact, the risk for cigarette use initiation peaks around age 16 and drops dramatically by age 20 (Chen & Kandel, 1985); however, recent evidence suggests that a sizable minority of college students initiate smoking after the age of 18 (Everett et al., 1999; Reed, Wang, Shillington, Clapp, & Lange, 2007). Earlier ages of onset for cigarette smoking is associated with a greater likelihood of nicotine addiction. Breslau & Peterson (1996) reported that those whose age of onset for cigarette smoking was age 13 or younger were significantly less likely to quit smoking compared to those who started at older ages. Further, unlike the use of many other substances that matures out -- that is the prevalence of heavy use decreases with increasing age -- cigarette smoking continues to be heavy and persistent throughout adulthood (Chen and Kandel, 1985). Half of all males who initiate smoking during adolescence can be expected to continue smoking until age 33 and half of females will continue smoking until age 37 (Pierce & Gilpin, 1996). Not only do young smokers continue to smoke for decades later, but they tend to develop into heavier smokers with age.

Given smoking begins at an early age and the possible “gateway” role tobacco plays in the progression of substance use behaviors, it is important to examine self-report stability as it relates to cigarette smoking behaviors. The onset of cigarette use and the age of first use are important characteristics in understanding the natural history of nicotine use and dependence. Furthermore, cigarette use is one of the earlier substance use behaviors often used to understand other substance use trajectories. Thus, for this study, we examined the report stability of cigarette ever use and smoking age of onset over ten years of data from the National Longitudinal Survey of Youth (NLSY). Internal consistency is also evaluated over six years of self-reports. The three types of report stability are examined across different demographic characteristics.

Methods

Sample

The NLSY is conducted in the U.S. and uses a multistage stratified random sampling technique. Originally 5,828 females and 5,578 males, aged 14-21 were interviewed with an over-sampling of Blacks and Hispanics. These same participants have been interviewed annually since 1979 with a retention rate at twelve year follow-up of 90.5% (Baker et al., 1993). In 1986 the study protocol was expanded to include the children of the female respondents. These children, however, were born to young mothers and therefore are not a nationally representative sample of children.

The children, when aged 10 and older, were asked to complete a self-report instrument entitled the Child Self Administered Supplement (CSAS). This

instrument queried areas such as child-parent relationships, attitudes, religiosity, deviance, and substance use. The interviews of the children have been conducted every two years. In 1994 a new instrument was added to the protocol. While children aged 10-14 still completed the CSAS, the new supplement was given to adolescents 15 years and older. The new survey was entitled the Young Adult Survey (YAS). The YAS asked about cigarette use, alcohol and other substance use, sexual activity, delinquent activities and relationships. The sample for this study consists of youth aged 10 and older beginning with those interviewed from 1988 to 1998.

Measures

For the purposes of this study, the first wave of data (Time 1) consisted of responses to cigarette use questions in the 1988, 1990, 1992, 1994, 1996 and 1998 surveys. Those who were in the younger age group responded to the CSAS while those who were 15 and older completed the YAS. Across the ten years, many children aged out of the CSAS and started responding to the YAS. Although the responses to the cigarette use questions were sometimes from two different survey instruments (CSAS and YAS), the questions that query substance use are nearly the same. For most waves of data a respondent would answer a question asking if they ever smoked a cigarette (no or yes). However there were 1-2 waves of data (depending on whether it was the CSAS or YAS) where the youth was asked to respond to a question regarding the number of times they'd smoked cigarettes. If a youth reported "0" cigarettes they were

coded as “no” for lifetime use and if he/she reported 1+ the code was a “1” for “yes” lifetime use.

For this study we used two variables to examine external consistency--agreement and discrepancy--for lifetime use of cigarettes from one wave of data to another. We also examined incident cases which are adolescents who report that they never used cigarettes at Time 1 but then did report use at Time 2. Because the two sources of data are from the adolescents' self-reports it is not possible to validate if new users from Time 1 to Time 2 are true incident cases. Thus incident cases are presented separately and not considered a type of report agreement.

Those reporting “never used” at both Time1 and Time 2 were excluded from our analyses for two reasons. First, prior research has shown that youth who report no use for each substance category are significantly younger than ever users (Shillington and Clapp, 2000). Second, the inclusion of never users in an analysis of report stability artificially decreases discrepancy rates (Bailey, et al., 1992).

External consistency

Users of a particular substance were coded as either “consistent” or “discrepant” for cigarette use for each wave of data to the next among those who reported cigarette use at a prior wave of data collection. Consistent reports for both Time 1 and Time 2 included adolescents reporting cigarette use at both waves of data collection. For example if a respondent reported lifetime use of

cigarettes at Time 1 and then reported such use again two years later they were categorized as in agreement or consistent in their reports for cigarette use.

Respondents could also be categorized as discrepant for cigarette use reports. Discrepant reports included those who reported use of a substance at Time 1 but two years later reported 'never used' cigarettes. For example, if an adolescent reported lifetime cigarette use during the 1992 survey but in 1994 responded that they never smoked a cigarette, then they met criteria for being discrepant in their reports. A variable was created for consistent use for six interviews across 5 timeframes 1988-90, 1990-92, 1992-94, 1994-96 and 1996-98.

Another aspect of report stability examined was the stability of the report of age at onset for cigarettes for each interview. A variable was created in which the difference was calculated between the age at onset reported at Time 2 and the age at onset reported at Time 1. This calculation could result in a "0" which would indicate that the age of onset was the same at both times. However, a value of "+3" for example would indicate that a youth reported their age of onset three years older at Time 2 compared to Time 1. A negative value would be for those reporting a younger age of onset from Time 2 compared to Time 1.

Internal consistency

To examine internal consistency of reports (consistency based on the logical skip pattern of the surveys), responses to two variables at each wave of data collection were compared. If the respondent reported lifetime use of cigarettes they were then asked to report how recently they smoked a cigarette.

If they reported smoking within the past 30 days they were asked to report the frequency of past 30 day smoking. Internally consistent reports were defined as those respondents who reported 30-day use and then reported they smoked cigarettes 1 or more times during the past 30 days. Inconsistent reports were defined as those respondents who reported past 30 day use but reported that they had “never smoked in the last 30 days” in response to the frequency of use question; thus, the survey respondents’ responses did not logically match. Inconsistent reports could also be defined as a situation where the youth reported no 30-day use but did answer the 30-day frequency question. These questions were available in the 1994, 1996 and 1998 interviews.

Demographic Variables

The dataset consisted of all youth who were age 10 or older by the last wave of data collection in 1998. Ethnicity was coded as 1= Hispanic, 2 = African American and 3 = Non-Hispanic/Non-African American. Age ranges changed based upon the year of interview. The minimum age was always 10 years but the upper age range increased with each interview as did the mean age. In 1988 the age range was 10 - 18 years with the upper age range increasing by two years at each later wave of data collection. By 1998 the mean age was 16.26 years with a range of 10 - 28 years. When age related analyses were conducted, only the chronological age at the time of the survey was used for the analyses. For example, when examining if there were any differences in report stability of cigarette use by chronological age, a separate analysis for each year’s stability variable and the same year’s age were conducted.

Statistical methods

Wave to wave comparison analyses were conducted using SPSS version 16.0. Chi-square analyses were conducted to compare the report agreement and discrepancy in reports by gender and ethnicity. Chronological age differences for those with report agreement and discrepancies was examined using analysis of variance and Scheffe's post-hoc means tests to reduce type I error. When testing for internal consistency with logically linked questions the analyses included a phi and Cramer's V coefficient. These analyses were conducted for each wave-to-wave comparison as well as the within year interviews for the internal consistency analyses.

Results

The total sample size of youth who were aged 10 years or older by the last interview in 1998 was 5,374. The sample was made up of 2,726 males (50.7%). The ethnic breakdown, as of 1998 for those 10 and older, was 1,191 Hispanic (22.2%), 1,897 African American (35.3%), and 2,286 Non-Hispanic/Non-African American (42.5%). The mean age of the sample changed for each wave of data collection as more children aged into the 10 and older group and were interviewed with the Child Self-Administered Supplement. Thus, the mean age for the 1,157 youth aged 10 and older in 1988 was 11.78 years while the mean age for the 5,374 youth aged 10 and older in 1998 was 16.26 years.

External consistency—Consistency of Use

As observed in Table 1, some change in the stability of self-reports was found from one wave of data to the next over the ten year period. From 1988-

1990 the agreement of use was 67.6% and was almost identical for the 1990 – 1992 interviews leaving about one third of respondents recanting their use at the wave two interview. Yet, as can be seen, the consistency of self-reported cigarette use went up to nearly 80% for the last six years analyzed with consistency ranging from 78.9% - 83.0%. Incident cases showed a steady decline over the ten years—starting at 54.9% from the 1988-1990 interviews and dropping to 32.3% by the 1996-1998 interviews.

TABLE 1 ABOUT HERE

In order to examine demographic differences for report stability, a dichotomous variable for report stability vs. recanting was created and chi-square analyses were run for the report stability variable with ethnicity and with gender. The results for ethnicity are presented in Table 2. The cell sizes were too small for the 1988-1990 analyses but the numbers were included for visual inspection. For the remaining four interview comparisons, significant differences were found. The lowest percentage of recanting was found for the Non-Hispanic, Non-African American youth with a recanting rate of 9.1% in 1994-1996 to a rate of 21.1% in 1992-1994. For Hispanic youth, recanting was observed for about one fifth of respondents. The lowest rate of recanting for Hispanics was 16.4% (1994-1996) but went up as high as 25.4% in the 1996-1998 waves of data. African-American respondents tended to have higher recanting rates, although the percentage of recanting among this group appears to decrease across time. In 1990 -1992 nearly half of African-American respondents recanted use at wave 2. However,

by the last waves of data (1996-1998), the recanting rate among African-American youth dropped to 27.5%, similar to that of Hispanic youth.

TABLE 2 ABOUT HERE

Similar to analyses for ethnicity, chi-square analyses were conducted for gender and report stability across the 10 years of data collection (not presented in a table). For 3 of the wave-to-wave comparisons, there were no statically significant differences found between males and females. But for two of the comparisons we found that males recanted more than females. From 1990-1992, 40.3% of males recanted compared to 23.6% of females ($X^2 = 4.78$, $p < .05$). The comparisons of reported cigarette use from 1992-1994 showed that 29.3% of males and 13.1% of females recanted their use ($X^2 = 12.25$, $p < .001$).

The last set of analyses for the comparisons of report stability over the ten year period of reports with chronological age. These data are shown in Table 3. The mean age for cigarette smokers is shown for the five wave-to-wave comparisons. The chronological mean age for those who recanted or denied their prior cigarette use was approximately six months of a full year younger than those who were consistent reporters across the 2 waves of data. The comparison across 1990-1992 was the only non-significant comparison. The mean age for consistent reporters ranged from 15.6 years to almost 18 years while the mean age for recanters was 14.4 years to 17.3 years.

TABLE 3 ABOUT HERE

External consistency—Age of Onset

The other part of external consistency, that is, consistency from one wave of data collection to the next is the reported age of onset. The data for the initial examination of the age of onset consistency is presented in Table 4. The first column in this table represents the percentage of cigarette users who reported their age of onset as the exact same age during each wave of data collection. The last three columns are the percentages who a) reported their age of onset within one year of the age reported at wave one, b) the percentage who reported their age of onset more than two years younger than at wave one (backward telescoping), and c) the percentage reporting the age of onset for cigarette use age two years or more older (forward telescoping) than the age reported at wave one. Thus, the last three columns add up to 100% of those who reported the age of onset differently at wave two than at wave one.

Table 4 ABOUT HERE

As can be seen in Table 4, across the ten years of data collection, the consistency of the reported age of onset for cigarette use remained below 30%. The lowest consistency was 15.4% in the 1988 – 1990 waves of data. There appears to be an increase in age of onset report consistency across the ten years with consistency rising to 27.8% in 1996-1998. When a one-year discrepancy allowance was made (i.e., the age of onset could be a year older OR a year younger from wave 1 to wave 2), the consistency increased substantially. Allowing for a one year differential increased consistency from 43.6% in 1988-1990 up to 60.0% in 1994-1996. Of those not captured by this one year allowance, the percentage of respondents who were two or more years **younger**

(backwards telescoping) in their reported age at wave 2 ranged from 6.2% to 15.4%. However, there appears to be a greater tendency for young cigarette smokers to report their age of onset as **older** (forward telescoping) at wave 2 than their stated age at wave 1. Two of the wave-by-wave comparisons (1988-1990 and 1992-1994) showed that nearly 40% of age of onset report inconsistency was among those reporting their age of onset as 2+years older at wave 2 relative to wave 1. The remaining three comparisons showed that about 1/3 of smokers reporting their age as older at wave 2 than wave 1.

In order to identify if there were any differences with gender or ethnicity for reported age of onset consistency, ANOVAs and Scheffe means tests were conducted. Results of the omnibus F-tests showed no significant differences for ethnicity or gender. We also calculated Pearson Correlations to examine if the consistency of reported age of onset across the two waves is associated with chronological age. Of the five wave-to-wave comparisons, only one was statistically significant (1992-1994) with a correlation of $-.23$ ($p < .02$), indicating that increasing chronological age was associated with stating a younger age of onset of cigarette use.

Internal consistency

Internal consistency is a within year comparison of the logical consistency of responses to questions regarding frequency and recency of cigarette use. These questions were asked from 1994-1998. Results from the Chi-square analyses are presented in Table 5. Results showed the inconsistency between reported recency of use and frequency of use for the past 30 days ranged from

9.4% in 1996 to 11.5% in 1994 with a Phi and Cramer's V tests significant for each of waves of data at $p < .001$. The majority of inconsistencies were for those who reported no use in the past 30 days but then answered the 30-day frequency question.

TABLE 5 ABOUT HERE

A comparison on ethnicity, gender and age were conducted to examine if differences existed for internal consistency (not in a table). The results for ethnicity indicate there were significant differences for two of the three years of data, 1994 and 1996. For those two years, internal consistency was high ranging from 84%-93. Specifically, Non-Hispanic, Non-African American group had the lowest report inconsistency (6.7 – 8.2%), followed by Hispanic youth (9.7 – 12.6%), and African Americans (13.8-15.9%).

Gender comparisons for internal consistency resulted in only one significant difference. In 1998, we found males had a significantly higher percentage of inconsistent internal reports compared to females (13.3% vs 8.4%; ($X^2 = 6.38, p < .01$). Males had higher percentages of inconsistent reports with the other two interviews but not significantly so. Lastly a comparison of internal consistency was made for the three years by chronological age. Results showed a significant correlation between age and internal consistency ($r = -0.09, p < 0.01$) indicating that with increasing age cigarette smokers were more consistent in answering the frequency and recency of smoking questions.

Discussion

External Consistency--Report Stability

Our analyses revealed changes over time for the percentage of youth who reported cigarette use at a prior wave of data collection but denied it two years later. During the first two wave comparisons we found about one third of all those reporting cigarette use at a prior wave of data collection denied or recanted it two years later. But for the latter years analyzed, the recanting rate decreased markedly. For the 1992-94, 1994-96 and 996-98 wave comparisons, we found recanting went down to about 20% or less.

Report stability across time was investigated for potential differences based upon respondent characteristics. These analyses revealed that differences, based upon ethnicity, were found for each wave comparison. Generally the Non-Hispanic/Non-African American group had the lowest discrepancy rates for wave-to-wave cigarette use reports followed by Hispanic teens. African-American youth had the highest rates of recanting. This is similar to prior research (Fendrich and Rosenbaum, 2003; Shillington and Clapp, 2001; Fendrich and Kim, 2001; Fendrich and Vaughn, 1994). Stanton et al., (2007) note that there may be cultural differences in the meaning of “ever smoked” for varying ethnic groups. Other investigators have noted that there may be less trust among ethnic minority groups of research in general. And finally there may be differences between groups in the perceived stigmatization of cigarette smoking.

Males have been found to have higher recanting rates (Stanton, Papandonatos, Lloyd-Richardson & Niaura, 2007; Percy, McAlister, Higgins, McCrystal & Thornton, 2005) but with using two waves of data. Fendrich &

Rosenbaum (2003) reported mixed differences for gender and recanting other substance use but did not identify any gender differences for cigarette use. The present study examined gender and recanting associations over ten years and six waves of data, and found that the gender association with recanting was not consistent across the varying waves of data.

Upon examination of report stability and chronological age, results showed that for all but one wave-to-wave comparison, those who denied or recanted their prior report of cigarette use were significantly younger than those who continued to report use. The age difference between the two groups was about nine to twelve months. Few studies have examined the association of chronological age with substance use recanting among adolescents. However, Shillington & Clapp (2000) reported recanting adolescents being younger than those who did not recant. This is in contrast to Stanton et al., (2007) findings that those who recant are older. One difference between the two studies is the amount of time between interviews. The Shillington & Clapp (2000) study examined data that were collected two years apart. For the Stanton (2007) study the data were collected with two surveys that were placed six years apart.

The cause of recanting remains unclear, although recanting has been identified as an international phenomenon. Percy, et al., (2005) reported recanting of previously reported drug use at a 1 year follow-up. Interestingly, they reported adolescent recanting increased after being exposed to an antidrug education program. Thus, some recanting may be to deny behaviors that have become perceived as socially undesirable. Other possibilities for recanting

include: the behavior reported at the first interview was untrue or exaggerated; the question was poorly understood by the respondent at either wave of data collection; or true forgetting. Fendrich & Kim (2001) further suggest recanting might be an artifact of repeated interviews found with longitudinal panel studies in that the youth may feel less anonymous and thus, social desirability is more salient. Despite the reasons for the recanting, this is the more serious of the report stability issues studied. Recanting a prior positive endorsement of lifetime use of a substance equates with denying an actual behavior. In contrast, report instability related to questions of age of onset are important errors, but because these types of questions ask respondents to report a higher degree of specificity, these errors may occur for different reasons (i.e., reasons not related to self-presentation or social desirability).

Epidemiological data is reliant upon self-reports of behavior. Any longitudinal data collection aimed at identifying trends in substance use over time would be impacted by how respondents answer an “ever use” question. A more costly approach is using Audio Computer Assisted Self Interviews (ACASI) which is what occurred in 2000 with the NLSY study. Using such technology allows for programming that skips all lifetime use questions if the respondent previously reported use in a prior survey. This solves the problem of recanting; however, it assumes the prior report of use was accurate. If recanting was due to exaggeration in prior behavior or misunderstanding or careless reading of the question, then those errors are carried forward. Therefore it is suggested that

the programming of the surveys not skip out of these questions even if a youth did report the use previously.

External Consistency—Age of Onset

To examine the consistency of self-reported age of onset for cigarette use, we compared the age of onset reported at wave one with that of wave two. The rates of those who reported their age of onset exactly the same from wave to wave was low with the highest rate being 30%. However the percentage of the smokers who reported their age of onset accurately from wave to wave increased from 1988 to 1998. We also allowed for a ± 1 year difference from wave one to wave two as per Wittchen, Burke, Semler, Pfiser, Vranach, and Zaudig (1989). This one year allowance consistently improved the percentage by 30%. With this one year allowance, the consistency of age of onset increased from about 44% up to 59% in 1998. We did find a higher percentage of those considered inconsistent reporters (those with 2+ years difference from wave one to wave two) to be forward telescopers rather than backward telescopers. About 10-15% of inconsistent age of onset reporters were backward telescopers. This means these respondents misreported their age of onset more than two years younger at wave two than at wave one. A larger percentage of inconsistent age of onset reporters were forward telescopers ranging from 31-41% of all inconsistencies. We did not find any meaningful differences in age of onset accuracy as a function of participant demographics. Our findings are similar to those reported by Johnson & Mott (2001) in which they found a rate of tobacco age of onset consistency of 54%. Also, our findings are similar to Golub, et al., (2000a) who

reported the average forward telescoping was about one year older for tobacco use among adolescents.

Most prior research on age of onset report consistency has focused on forward telescoping. As Johnson & Schultz (2005) point out, descriptions of the course of a disorder and the risks that are associated with it when onset occurs may be inaccurate. Forward telescoping was found previously for both experimental and daily smokers (Johnson & Schultz, 2005). Forward telescoping is the type of discrepancy focused on the most because the majority of those with discrepancies are in that direction. It is unclear if the discrepancy is intentional or results from a change in the respondent's interpretation of the survey question. Such changes across time could indicate a respondent perceives a later age is more appropriate or more normative for the age of first use. In terms of interpreting the question differently, when they are asked at an earlier time if they ever smoked, they may give an age when they had a couple puffs from a cigarette. But at the later interview, two years subsequent, their interpretation of the question is the age when they began smoking cigarettes regularly. Despite the reason why it occurs, these discrepancies are important for researchers who studying substance use progression and trajectories.

4.3. Internal consistency

The last part of report consistency that was examined in this study was logical consistency within three waves of data. This was a logical response to one question regarding recency of cigarette use and subsequent questions concerning past-30 day frequency of use. Logical consistency was high, at about

90%. When the inconsistencies were examined by respondent characteristics, we found higher inconsistencies for ethnic minorities, males and younger respondents but these differences were not found in each of the survey years.

Strengths and Limitations

This is the first paper to systematically examine wave by wave report stability for self-reports of smoking behavior for a ten year period. The panel design of the NLSY is both a strength and a limitation to the present study. The large and ethnically diverse sample of youth, coupled with a very low attrition rate, allow for detailed wave-to-wave analyses of specific ethnic groups over long periods. However, these data have limited external validity as the adolescents were born to younger mothers.

All secondary analyses are limited by the investigators' inability to alter survey measures or study procedures. The NLSY was not designed explicitly to measure substance use. Nor was the survey designed with the intent to assess reliability issues.

Implications and Future Directions

Better measures of substance use, including timeline follow-back for recent use and clinical measures would have enhanced the present study. Additionally, better recall prompts would benefit future studies. More research is needed to understand the relatively consistent finding that ethnic minority youth are more likely to recant prior reports of use relative to their non-minority peers. More needs to be understood regarding if the recanting is due to over-reporting at time one, recanting at time two, and if it is recanting, is it due to a fear of

stigmatization or other consequences. Many large studies have started using ACASI. This technology will help in data consistency. So if a person reported use at time one the same questions are skipped thereafter in later waves of data collection. However, such technology will not help untangle why such group differences exist in recanting.

The present study was unable to examine cognitive reasons related to reporting errors. Such work is greatly needed in future studies. Given that intervention and prevention programs are often grounded in findings that were derived from self-reports, understanding the stability of these reports is essential.

References

- Andreasen, N. C., Grove, W. M., Shapiro, R. W., Keller, M. B., Hirschfeld, R. M. A., & McDonald-Scott, P. (1981) Reliability of lifetime diagnosis: a multi-center collaborative perspective. *Arch Gen Psychiat*, 38, 400-405.
- Bailey, S. L., Flewelling, R.L., and Rachal, V.J. 1992. The characterization of inconsistencies in self-reports of alcohol and marijuana use in a longitudinal study of adolescents. *J. Stud. Alc.*, 53, 636-647.
- Baker, P. C, Keck, C. K, Mott, F.L., & Quinlan, S.V. (1993). NLSY Child Handbook- Revised Edition: A Guide to the 1986-1990 National Longitudinal Survey of Youth Child Data. Columbus, OH: Center for Human Resource Research, October.
- Barnea, Z, Rahav, G & Teichman, M. (1987). The reliability and consistency of self-reports on substance use in a longitudinal study. *British J of Addiction*, 82, 891-898.
- CDC (2005). Cigarette smoking among adults—United States, 2004. *Morbidity and Mortality Weekly Report*, 54, 1121-1124.
- CDC (2006). Cigarette use among high school students—United States, 1991-2005. *Morbidity and Mortality Weekly Report*, 55, 724-726.
- Chen, K., & Kandel, D.B. (1985) The natural history of drug use from adolescence to the mid-thirties in a general population sample. *Am J Public Health*. 85, 41-47.
- Everett, S. A., Husten, C. G., Kann, L., Warren, C.W., Sharp, D., & Crossett, L. (1999). Smoking initiation and smoking patterns among US college students. *Journal of American College Health*, 48, 55–60.
- Fendrich, M., Weissman, M. M., Warner, V., & Mufson, L. (1990) Two-year recall of lifetime diagnoses in offspring at high and low risk for major depression. *Arch Gen Psychiat*, 47, 1121-1127.
- Fendrich, M & Vaughn, CM (1994). Diminished lifetime substance use over time. *Pub Opin Quarterly*, 58, 96-123.
- Fendrich, M. & Mackesy-Amiti, ME. (1995). Inconsistencies in lifetime cocaine and marijuana use reports: impact on prevalence and incidence. *Addiction*, 90, 111-118.

- Fendrich, M & Kim, J.Y.S. (2001). Multiwave analysis of retest artifact in the National Longitudinal Survey of Youth drug use. *Drug and Alc Dep*, 62, 239-253. PII: S0376-8716(00)00177-0
- Fendrich, M & Rosenbaum, DP. (2003). Recanting of substance use reports in a longitudinal prevention study. *Drug and Alc Dep*, 70, 241-253. doi:10.1016/S0376-8716(03)00010-3
- Funkhouser, JE, Goplerud, EN & Bass, RO. (1992). The current status of prevention strategies. In Jansen, MA (ed.), *A Promising Future: Alcohol and Other Drug Problem Prevention Services Improvement*, OSAP Prevention Monograph No. 10, DHHS) Pub No. (ADM) 92-1807, Gov Printing Office, Washington, DC, pp 17-82.
- Grant, BF, (1998). Age at smoking onset and its association with alcohol consumption and dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. *J Sub Abuse*, 10, 59-73.
- Grant, BF, & Dawson, DA. (1997). Age at onset to alcohol use and its association with DSM-IV alcohol abuse and dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. *J Sub Abuse*, 9, 103-110.
- Golub, A, Johnson, BD, & Labouvie,E. (2000a). On correcting biases in self-reports of age at first substance use with repeated cross-section analysis. *J Quant Criminology*. 16, 45-67.
- Golub, A, Labouvie,E. & Johnson, BD, (2000b). Response reliability and the study of adolescent substance use progression. *J of Drug Issues*. 30, 103-118.
- Hawkins, B., Catalano, RP & Miller JY (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychol Bull*, 112, 64-105.
- Johnson, RA, Gerstein, DR, & Rasinski, KA (1997). *Recall decay and telescoping in self-reports of alcohol and marijuana use: Results from the National Household Survey on Drug Abuse (NHSDA)*. In Proc. Am. Assoc. Pub. Opin. Research.
- Johnson, TP & Mott, JA. (2001). The reliability of self-reported age of onset of tobacco, alcohol and illicit drug use. *Addiction*, 96, 1187-1198. DOI: 10.1080/09652140120060770
- Johnson, EO & Schultz L. (2005). Forward telescoping bias in reported age of onset: an example from cigarette smoking. *Internat'l J of Methods in Psychiatric Research*. 14(3), 119-129. DOI: 10.1002/mpr.2

- Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2008). Monitoring the Future national results on adolescent drug use: Overview of key findings, 2007 (NIH Publication No. 08-6418). Bethesda, MD: National Institute on Drug Abuse, 70 pp.
<http://monitoringthefuture.org/pubs.html> accessed 4.21.08.
- Kandel, DB, Yamaguchi, K, & Chen, K. (1992). Stages of progression in drug involvement from adolescence to adulthood: Further evidence for the gateway theory. *J Studies on Alc*, 53(5), 447-457.
- Mills, CJ & Noyes, HL. (1984). Patterns and correlates of initial and subsequent drug use among adolescent. *J Consul Clin Psychol*, 52, 231-243.
- Mosbach, P. & Leventhal, H. (1988). Peer group identification and smoking: implications for intervention. *J Abnorm Psycho*, 97, 283-45.
- NIDA, Monitoring the Future Study: Trends in prevalence of various drugs for 8th-graders, 10th-graders, and high school seniors. (2007). Institute for Social Research. University of Michigan for the National Institute on Drug Abuse, National Institutes of Health.
- O'Malley, P.M., Bachman, J.G., & Johnston, L. D. (1983). Reliability and consistency in self-reports of drug use. *Internat. J. Add.*, 18, 805-824.
- Pedhazur, EJ, & Schmelkin, LP. (1991). *Measurement, Design and Analysis: An Integrated Approach*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Percy A, McAlister, S, Higgins, K, McCrystal,P, & Thornton, M. (2005). Response consistency in young adolescents' drug use self-reports: a recanting rate analysis. *Addiction*, 100, 189-196. doi:10.1111/j.1360-0443.2004.00943.x
- Peto, R., Lopez, A., Boreham, J. et al. (1994). *Mortality from Smoking in Developed Countries*. New York: Oxford University Press.
- Pierce, J.P. & Gilpin, M.S. (1996). How long will today's new adolescent smoker be addicted to cigarettes? *Am J Public Health*, 86, 253-256.
- Prause J, Dooley, D, Ham-Rowbottom, KA, & Emptage, N. (2007). Alcohol drinking onset: a reliability study. *J of Child & Adolescent Sub Abuse*, 16, 79-90. doi: 10.1300/J029v16n04_05

- Prusoff, B.A., Merikangas, K.R., & Weissman, M.M. (1988). Lifetime prevalence and age of onset of psychiatric disorders: recall 4 years later. *J. Psych. Res.* 22, 107-117.
- Reed, M.B., Wang, R., Shillington, A.M., Clapp, J.D., & Lange, J.E. (2007). The relationship between alcohol use and cigarette smoking in a sample of undergraduate college students. *Addictive Behaviors*, 32, 449-464. <http://dx.doi.org/10.1016/j.addbeh.2006.05.016>
- Rice, J. P., Rochberg, N., Endicott, J., Lavori, P. W., & Miller, C. (1992) Stability of psychiatric Diagnoses: An application to the affective disorders. *Arch Gen Psychiat.* 49, 824-830.
- Rosen, I.M. & Maurer, D.M. (2008). Reducing tobacco use in adolescents. *American Family Physician*, 77, 483-490.
- Smith, G.T., McCarthy, D. M. & Goldman, M.S. (1995). Self-reported drinking and alcohol-related problems among early adolescents: Dimensionality and validity over 24 months. *J Stud Alc*, 56, 383-394.
- Shillington, AM, Cottler, LB, Mager, D.E., & Compton, W.M. 1995. Self-report stability for substance use over 10 years: Data from the St. Louis Epidemiological Catchment Study. *Drug Alc. Dep.*, 40, 103-109.
- Shillington, AM & Clapp, JD. Self-Report Stability of Adolescent Substance Use: A First Look at Gender, Ethnicity, and Behavioral Problems. *Drug and Alcohol Dependence. An International Journal on Biomedical and Psychosocial Approaches.* 60, 1, 19-27, 2000. PII: S0376-8716(99)00137-4
- Siddiqui, O, Mott, JA, Anderson, TL, & Flay, Br. (1999) Characteristics of inconsistent respondents who have "ever used" drugs in a school-based sample. *Subt. Use Misuse* 34, 269-295.
- Stanton, CA, Papandonatos, G, Lloyd-Richardson, EE, & Niaura, R. (2007). Consistency of self-reported smoking over a 6-year interval from adolescence to young adulthood. *Addiction*, 102, 1831-1839. doi:10.1111/j.1360-0443.2007.01974.x
- Wittchen, H. U., Burke, J. D., Semler, G., Pfister, H., Cranach, V., & Zaudig, M. (1989) Recall and dating of psychiatric symptoms: Test-retest reliability of time-related symptom questions in a standardized psychiatric interview. *Arch Gen Psychiat* 46, 437-443.

Table 1. Report agreement, report discrepancy and incident cases for users over ten years.

Years of Interviews	Total N Reported Use Time 1 or Time 2	Agreement Yes-Yes N (%)	Use Discrepancy Yes- No N (%)	Incident Cases N (%)
1988-90	144	44 (67.6%)	21 (32.4%)	79 (54.9%)
1990-92	292	101 (67.8%)	48 (32.2%)	143 (48.9%)
1992-94	542	239 (78.9%)	64 (21.2%)	239 (44.1%)
1994-96	891	414 (83.0%)	85 (17.0%)	392 (43.0%)
1996-98	1014	547 (79.7%)	139 (20.3%)	328 (32.3%)

Table 2. Report agreement and report discrepancy by ethnicity for users over ten years.

	Agreement Yes-Yes N (%)	Use Discrepancy Yes- No N (%)	Chi-Square Value ^a
1988-1990			
Hispanic	5 (55.6)	4 (44.4)	
African American	18 (58.1)	13 (41.9)	
Non-Hispanic, Non-African American	21 (84.0)	4 (16.0)	Cell sizes too small
1990-1992			
Hispanic	20 (76.9)	6 (23.1)	
African American	34 (50.7)	33 (49.3)	
Non-Hispanic, Non-African American	47 (83.9)	9 (16.1)	16.9***
1992-1994			
Hispanic	50 (76.9)	15 (23.1)	
African American	59 (63.4)	34 (36.6)	
Non-Hispanic, Non-African American	130 (89.7)	64 (21.1)	23.64***
1994-1996			
Hispanic	97 (83.6)	19 (16.4)	
African American	97 (68.8)	44 (31.2)	
Non-Hispanic, Non-African American	220 (90.9)	22 (9.1)	29.57***
1996-1998			
Hispanic	135 (74.6)	46 (25.4)	
African American	124 (72.5)	47 (27.5)	
Non-Hispanic, Non-African American	288 (86.2)	46 (13.8)	17.47***

^a Fisher's Exact Test

* p < 0.05, ** p<.001, *** p<.0

Table 3. Means ANOVA test results for chronological age and report consistency or inconsistency for cigarette use from Wave 1 to Wave 2.

VARIABLE	Report Cigarette Use Both Waves Mean Age (SD)	Denied Use at Wave 2 Mean Age (SD)	ANOVA Value
1988 - 1990	15.60 (1.70)	14.43 (1.89)	6.18*
1990 - 1992	15.68 (2.03)	15.46 (2.03)	ns
1992 - 1994	16.78 (2.13)	16.10 (2.22)	4.81*
1994 - 1996	18.00 (2.36)	17.08 (2.91)	9.53**
1996 - 1998	17.95 (1.96)	17.32 (2.42)	10.43***

ns = not statistically significant

* = $p < .05$

** = $p < .01$

*** = $p < .001$

Table 4. Percentage who reported their age of onset accurately and the variability of inaccuracy across waves of data collection.

Reported Age of Onset Exactly the Same Wave to Wave	Years of Interviews	± 1 YEAR	Backward Telescopers ≤ 2 YEARS	Forward Telescopers ≥ 2 YEARS
		Reported Age of onset 1 Year Older or Younger	Reported Age of onset 2 or more Years YOUNGER at Wave 2	Reported Age of Onset 2 or more Years OLDER at Wave 2
15.4%	1988-90	43.6%	15.4%	41.0%
22.2%	1990-92	55.5%	12.1%	32.4%
24.2%	1992-94	54.6%	6.2%	39.2%
26.1%	1994-96	60.0%	6.6%	33.4%
27.8%	1996-98	59.1%	9.3%	31.6%

Table 5. Internal consistency between the recency of use question and the frequency of use question.

Substance	Recency of last use		Total Inconsistencies	Phi and Cramer's V
	Used past 30 days n (%)	No use past 30 days n (%)		
1994				
No use past 30 days	15 (2.5)	219 (48.7)	69 (11.5)	.775***
Used 1+ times past 30 days	238 (39.8)	54 (9.0)		
1996				
No use past 30 days	10 (1.1)	386 (43.1)	84 (9.4)	.821***
Used 1+ times past 30 days	426 (47.5)	74 (8.3)		
1998				
No use past 30 days	11 (1.1)	420 (40.6)	87 (10.2)	.798***
Used 1+ times past 30 days	504 (48.7)	100 (9.7)		

ns = not statistically significant

* = p < .05

** = p < .01

*** = p < .001

Cigarette Report Stability Over 10 Years