A RAND NOTE

How Accurate Are Adolescent Reports of Drug Use?

Ellen J. Reinisch, Robert M. Bell,
Phyllis L. Ellickson

RAND

The research described in this report was supported by a grant from the Conrad N. Hilton Foundation.

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of RAND do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published 1991 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

A RAND NOTE

N-3189-CHF

How Accurate Are Adolescent Reports of Drug Use?

Ellen J. Reinisch, Robert M. Bell,
Phyllis L. Ellickson

Supported by the
Conrad N. Hilton Foundation

RAND

PREFACE

The Project ALERT experiment and associated research were supported by a grant from the Conrad N. Hilton Foundation. The study this Note reports was undertaken to assess the implications of inconsistent or incomplete responses regarding the accuracy of self-reported substance-use data. The findings should interest researchers concerned with the validity of data obtained through self-reports.

SUMMARY

Project ALERT (Adolescent Learning Experiences in Resistance Training), funded by the Conrad N. Hilton Foundation, is a multiwave study that tests the effect of a drug-prevention program for seventh- and eighth-grade students. Because it relies on self-reports to obtain information on whether and how often students use specific substances, assessing whether the collected data accurately represent the respondents' use is important. We were able to use laboratory results of the students' saliva to confirm self-reports concerning recent cigarette use but had no such form of external validation for alcohol, marijuana, or less recent cigarette use. Thus, we examined the consistency and completeness of student responses within and across questionnaires to gain information about the reports' accuracy. In addition to determining how often inconsistent or incomplete data occurred, we assessed how the problems arose and what impact they might have on data bias and treatment effects analyses.

This Note covers the program's first four waves of data. Wave 1 took place during the seventh grade, before the treatment groups received their first curriculum session. The second wave followed three months later. Waves 3 and 4 took place in the eighth grade, approximately 12 and 15 months after baseline. The questionnaire concentrated on three "target" substances: cigarettes, alcohol, and marijuana. Respondents were asked a battery of questions about their experience with each substance—recency of use, frequency, level or dosage, and indicators of abusive use. We focused on this battery to identify incomplete or inconsistent data both within and across questionnaires. We examined three types of longitudinal inconsistencies:

- **Retractions.** A retraction occurred at a particular wave when a student denied any lifetime use after having admitted to use at an earlier wave.
- **Modifications.** A modification occurred when a student denied use in the past year after admitting to use in the past month at the previous wave (three to nine months earlier).
- **New admissions.** A new admission occurred when a student admitted to use but not within the past year, after having denied use at all previous waves.

Less than 5 percent of the students provided incomplete or inconsistent responses within questionnaires. Most discrepancies were attributable to unintentional errors caused by carelessness or misunderstanding rather than deliberate distortions.

More than 40 percent of the students committed at least one longitudinal inconsistency across the four waves of data collection. However, over 95 percent of these errors were minor—classifiable as either inconsistent reporting of infrequent experimental use or as inaccurate recall of whether limited prior use had occurred within the past year. Severe inconsistencies, which involved denying frequent use, were committed by less than 2 percent of the students.

Laboratory assessments of saliva provided independent verification that the majority of students accurately reported recent cigarette use. Of students identified as recent tobacco users based on the laboratory tests, 95 percent reported using cigarettes or chewing tobacco in the past month. Because the percentages of retractions, modifications, and new admissions were in the same range for all three substances, we conclude that the accuracy level for alcohol and marijuana was similar to that for cigarettes.

Reporting errors do not threaten the validity of the treatment effects analysis. Inconsistencies involving initiation of use or denial of frequent or recent use would be cause for concern if the percentages varied by treatment group. Because we found little evidence of significant differences, any inaccuracies should be independent of group and thus in no way bias treatment effects.

We conclude that the overwhelming majority of Project ALERT students accurately reported their drug use. By identifying and examining the incomplete and inconsistent cases, we have also eliminated the few students with completely invalid data from the analyses.

ACKNOWLEDGMENTS

We are grateful to the Conrad N. Hilton Foundation for its commitment to drug prevention and its support of this research. There were many people involved in creating the surveys, implementing data collection, and constructing analysis files. We acknowledge their efforts and thank them for their hard work. In addition, we appreciate the cooperation of the students who participated in the study. Finally, we thank Susan Marquis for her helpful review.

CONTENTS

FIGURES

TABLES

I. INTRODUCTION

Over the past ten years, drug use among adolescents has become a growing concern; consequently, more and more studies have sought to determine its prevalence among young teens and to test new techniques designed to prevent or diminish drug use. Among these studies is Project ALERT (Adolescent Learning Experiences in Resistance Training), a multiwave study that tests the effects of a drug-prevention program for seventh- and eighth-grade students. Like other drug-use studies, this study relies on self-reports to obtain information on whether and how often students use specific substances. Thus, before we address prevalence or treatment effect issues, we must assess whether the collected data are true representations of the respondents' use.

PROCEDURES TO IMPROVE THE ACCURACY OF DRUG-USE SELF-REPORTS

Data reported elsewhere suggest that most students are willing to provide self-reports and give accurate accounts of recent tobacco use (Barnea, Rahav, and Teichman, 1987; Mensch and Kandel, 1988; O'Malley, Bachman, and Johnston, 1983; Single, Kandel, and Johnson, 1975). Research has also shown that taking certain steps during the design and implementation of data collection can increase the likelihood of obtaining accurate data (Gfroerer, 1985; Johnston and O'Malley, 1985). In designing Project ALERT, we instituted several such procedures. To reduce incomplete or inaccurate responses, the questions had simple, explicit response options; we then pilot-tested them to identify any possible comprehension problems. To encourage truthfulness, we guaranteed the privacy and confidentiality of the data, used data collectors whom the students did not know, and collected saliva samples from each participant immediately before survey administration. We also gave students the opportunity to refuse participation.

Fewer than 1 percent of the baseline sample refused to participate, indicating that students were at least willing to answer the questionnaires. Because 95 percent of students identified as recent tobacco users based on laboratory assessments of their saliva had reported using cigarettes or chewing tobacco in the past month, we surmise that, at least for tobacco use, the students did not fear admitting current use. Unfortunately, we could not use the physiological data to confirm alcohol or marijuana self-reports—or survey data on less recent cigarette use. To gain information about these reports' accuracy, we examined the consistency and completeness of student responses within and across questionnaires.

POTENTIAL PROBLEMS WITH SELF-REPORTS

In addition to determining how often incomplete data or inconsistencies occurred, we assessed how these problems arose—whether they reflected intentional efforts to distort the truth or unintentional errors attributable to difficulties students had completing the questionnaire. Each situation is likely to occur, but for different reasons. Students might intentionally distort the truth because substance use is socially disapproved (and/or illegal). Alternatively, unintentional discrepancies might occur because students find reading or understanding the questionnaire difficult, are careless, or have problems recalling their use over the entire period in question.

Each situation also has different implications for the expected effect on data bias. Unintentional problems, whether caused by comprehension difficulties or carelessness, should be independent of experimental group. Though they may affect the accuracy of prevalence estimates, they should not bias treatment effects analyses. Reasons for intentionally misrepresenting use may vary by wave. Students might conceal their use at baseline because they question the study's confidentiality, causing us to underestimate prevalence rates. However, such concealment should not vary across treatment and control groups and thus should not bias estimates of treatment effects. A more dangerous problem could occur at subsequent waves if students in the treatment group underreport their usage to give the appearance that the program has been successful. If the data's accuracy were related to experimental group, we would be unable to make a valid assessment of the treatment effects—the study's main objective.

GOAL OF THIS NOTE

In summary, this Note's goal is to determine the quality of the self-reported substance-use data and to assess the implications on subsequent analyses. We will do this by answering the following questions:

- How often do students give incomplete or inconsistent responses within a single questionnaire?
- How often do students give inconsistent responses across time?
- Do discrepant answers typically reflect deliberate or unintentional misreporting?

- Does the accuracy rate for self-reports vary by substance?
- How do the inconsistencies affect estimates of treatment effects?
- How does the accuracy of self-reported drug use among Project ALERT students compare with the results of other studies?

Section II presents the Project ALERT program, focusing on the data collection procedures and outcome measures. Section III presents the methods we used to evaluate the inconsistencies, first cross-sectionally and then longitudinally. In Sec. IV, we address the above questions, presenting our findings on the prevalence of inconsistencies, the certainty with which they can be classified, and what they reveal about the data's accuracy. Section V presents a summary of our conclusions and a final assessment of the data's accuracy.

II. DATA COLLECTION

We tested Project ALERT in 30 junior high schools from eight school districts throughout California and Oregon. We divided the treatment group into two arms. Ten schools received the Project ALERT curriculum from a health educator alone, and ten schools had an older teen leader assist the health educator. The remaining ten schools served as the control group and received no special curriculum. To prevent possible confounding of effects caused by students in treatment classes interacting with students in control classes, each class within a given school received the same treatment.

Although we collected a total of six waves of data, this Note focuses on the first two years (or four waves) of the program (see Fig. 2.1). Baseline (wave 1) took place during the seventh grade, before the treatment groups received their first curriculum session. The second wave followed three months later, approximately four weeks after the final curriculum session. Waves 3 and 4 took place in the eighth grade, approximately 12 and 15 months after baseline. Students in the treatment groups received three booster sessions following the third wave. The fifth and sixth waves took place in the ninth and tenth grades, respectively.

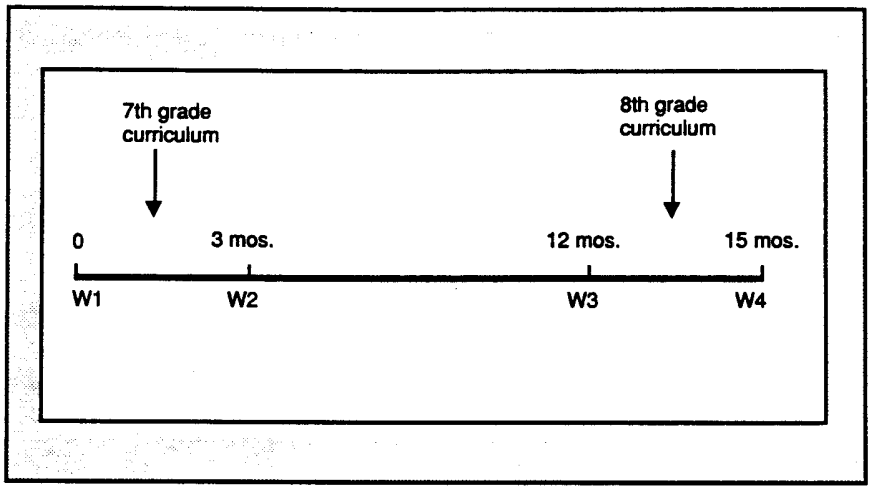

Fig. 2.1—Data collection and program delivery schedule

STUDY SAMPLE

We collected data on 6527 students at baseline. This Note focuses on the subset of 4605 students who completed a survey at all four waves. Admittedly, this subgroup may represent a biased sample in that students who refused to fill out a survey or moved were more likely to have used drugs—and perhaps more likely to provide inaccurate data. We rejected varying the sample based on the specific wave(s) of interest because we wanted to maintain the same students across analyses. If instead, we had included all baseline students, then analysis of longitudinal inconsistencies occurring at wave 3 would have included students who were not present at wave 2, and thus had much less opportunity to contradict previous answers. Thus, even though the larger sample might include a greater percentage of inconsistent students, we might detect a smaller percentage of inconsistent responses. To understand the effect altering the sample has on the apparent prevalence of inconsistencies, we examined inconsistencies at the first two or three waves relative to attendance at the subsequent wave(s). The findings, which are summarized in the appendix, show that those who did not attend the later waves were more likely to have been inconsistent at earlier waves, and that inconsistent students were slightly more likely to miss subsequent waves.

ENHANCING DATA QUALITY

In designing the questionnaire and data collection procedures, we sought to minimize incomplete or inaccurate responses to the drug-use questions. As Gfroerer (1985) reported, "the type of questionnaire and the degree of anonymity of the respondent are among the factors which influence the self-reporting of drug use by students."

We pilot-tested various versions of the questionnaire to identify possible problems with wording or instructions. To improve response accuracy, we used objective, explicit anchors on the response scales. Thus, responses indicated the number of times the students used drugs in specific time periods, rather than vague measures like "sometimes" or "often." Inconsistent interpretations of subjective choices would have reduced the comparability of responses across both students and waves.

To minimize intentional concealment (and intentional bragging), we paid special attention to guaranteeing the data's privacy and confidentiality. We assured students that none of their teachers, principals, or parents would see their responses; we used numbers rather than names as identifiers; and we distributed the questionnaires in a group setting rather than in a face-to-face interview. These techniques have been shown to enhance students' confidence that their answers will remain anonymous (Johnston and O'Malley,

1985; Mensch and Kandel, 1988). As a final measure to ensure the data's privacy, we obtained a certificate of confidentiality from the Department of Health and Human Services. Students also had the option of refusing to participate at any time. Because fewer than 1 percent of students refused to complete a survey at baseline, and only slightly more declined to participate at the subsequent waves, we believe that most students felt they could give honest answers without fear of retribution.

As other studies have shown, the mere act of collecting saliva samples from students improves the accuracy of self-reports (Bauman and Dent, 1982). Before distributing the questionnaires at each wave, we explained to the students that recent tobacco use was detectable in saliva. We then asked them to provide us with samples. This exercise served both physiological and psychological purposes. Radioimmunoassays to detect cotinine, a by-product of nicotine, were performed at a laboratory.[1] Though in theory, cotinine can be detected in the saliva of those who have chewed tobacco or smoked cigarettes up to three days earlier, the test is not infallible: Students who use a relatively small dosage may go undetected. Because the students did not know this, and believed that all users would be detected by their saliva, the saliva specimens also provided a further incentive to respond honestly to the questionnaire. Of the students that the laboratory test could identify as recent smokers, 95 percent admitted to using cigarettes or chewing tobacco in the past month.[2]

SUBSTANCE-USE DATA

The curriculum and questionnaire focused on three "target" substances: cigarettes, alcohol, and marijuana. It also included a limited number of questions that dealt with chewing tobacco and harder drugs such as cocaine, "uppers," and "downers." We asked several questions for each target substance, soliciting information about most recent use, frequency of use, dosage, and indicators of abusive use (Fig. 2.2). At the first two waves of data collection, we also asked the age at which students first used the three main substances. In general, we repeated the same substance use questions over the four waves so we could observe patterns of use such as starting, quitting, or changing the frequency of use. These repeated questions also provided the basis for identifying students who may have lied or misreported use at one or more waves.

---

[1]Because of a very narrow detection window, we did not use the saliva to identify recent marijuana use.

[2]To avoid identifying nonusing adolescents who live with smokers as smokers themselves, we used a cutoff point (10 nanograms per milliliter) that excludes cotinine levels potentially attributed to persons contaminated by passive smoke.

9. Have you EVER smoked a cigarette — even just a few puffs?

Yes	1 ---> **(Answer Questions 10-16)**
No	2 ---> **(Skip to Question 17)**

10. How many times have you smoked a cigarette in the *LAST YEAR*?

None	1
1 or 2 times	2
3 to 10 times	3
11 to 20 times	4
More than 20 times	5

11. On how many DAYS did you smoke a cigarette in the *LAST MONTH*?

None	1
1 or 2 days in the last month	2
3 to 5 days in the last month	3
6 to 19 days in the last month	4
20 or more days in the last month	5

12. Have you smoked a cigarette in the *LAST WEEK*?

Yes	1
No	2

13. Have you smoked a cigarette in the *LAST TWO DAYS*?[a]

Yes	1
No	2

14. On the days you smoke cigarettes, how many do you usually smoke?

Less than 1 cigarette a day	1
1 or 2 cigarettes a day	2
3 to 7 cigarettes a day	3
About 1/2 a pack of cigarettes a day	4
A pack or more of cigarettes a day	5
I don't smoke cigarettes	6

15. Do you ever smoke cigarettes when you're by yourself?

Yes, often	1
Yes, sometimes	2
No, never	3

16. Which of these statements best describes you *now*?

I smoke cigarettes often	1
I smoke cigarettes sometimes	2
I tried cigarettes once or twice	3
I used to smoke but I quit	4

70. How old were you when you *first* smoked a cigarette?

[a]Included only in cigarette battery

Fig. 2.2—Cigarette battery of questions (waves 1 and 2)

The surveys at the first two waves of data collection contained skip patterns for the recency-of-use questions (see Fig. 2.2). At baseline, approximately 10 percent of students who stated that they had never used a substance ignored the skip instruction and answered the remaining questions. Of greater concern, about 1 percent of students who admitted to lifetime use skipped the remaining questions. To reduce the potential for inadvertent inconsistencies caused by misreading the instructions, we dropped the skip pattern from subsequent waves. We also rearranged the items to ask about more recent periods first.

RECENCY OF USE

When students gave complete, consistent information for a substance, we were able to classify their use into a recency-of-use variable with the following values:

0 = never used;

1 = used in lifetime, but not in past year;

2 = used in past year, but not in past month;

3 = used in past month, but not in past week;

4 = used in past week (but not in past two days for cigarettes);

5 = used in past two days (for cigarette use only).

This variable was the key component in identifying both cross-sectional and longitudinal problems involving the target substances. Within a wave, incomplete or inconsistent responses to the battery of questions prevented us from immediately determining the recency of use. Once we were able to classify the recency-of-use, we compared the values across waves to identify longitudinal inconsistencies. Section III presents the methods we used to assess these inconsistencies.

III. METHODS

RELATIONSHIP BETWEEN INCONSISTENCIES AND ACCURACY

Do inconsistent or incomplete responses invalidate only the questions involved, the entire survey, or all waves of data collected for that student? Or do they indicate nothing at all? Certainly, one specific rule would not be appropriate for every situation. Though a student may provide consistent reports that are not accurate reflections of his actual use, inconsistent responses, deliberate or accidental, act as flags cautioning the analyst to examine the data more carefully. By checking the inconsistent data and other related questions, we can then assess the responses' accuracy. If the problem is caused by an unintentional response to an isolated question, enough information may have been gathered from surrounding consistent responses to infer with high confidence what the correct pattern should be. Thus to have entirely consistent responses with which to extract accurate information from self-reports is not always necessary. Of course, a few students will always choose to give inaccurate accounts of their use, whether by exaggeration or denial. Students who have not yet mastered the skill of deception will be detectable by inconsistent responses across the different waves. The frequency of this problem provides a lower-bound estimate of the percentage of self-reports of drug use by students that are inaccurate representations of the actual use.

To assess the seriousness of inconsistent data, we first attempted to identify the probable causes of the discrepancies. The reasons for incomplete or inconsistent data typically fall into four general categories: misunderstanding, poor recall, carelessness, and intentional misrepresentation. If most problems are caused by misunderstanding or poor recall, we can feel confident that the majority of students are attempting to provide honest reports, even though some may not be completely accurate. However, blatant inconsistencies, which are most likely attributable to deliberate or careless errors, give cause for serious doubt about the data's accuracy.

The remainder of this section presents the methods we used to assess the accuracy of the Project ALERT data. We first define the types of discrepancies that arose both within and across questionnaires. We then use additional survey information to better understand their causes and consequences. Finally, we look at the prevalence of inconsistent or incomplete data across different demographic characteristics to learn if the potential for inaccurate data varies by gender, ethnicity, or academic performance.

CROSS-SECTIONAL PROBLEMS

Incomplete Responses

Missing data could result from inadvertent omissions or the student's intentional decision to refuse to answer a question. The latter reason is more likely to be associated with use, while the former is likely to occur at random. In trying to understand why missing data occurred, we compared the frequency of incomplete response patterns across substances and waves. To assess whether missing data are a function of the question's sensitivity, we also looked at the frequency of incomplete responses for substance-use questions relative to missing data for less-threatening questions.

Inconsistent Responses

Inconsistent responses within a single wave arise when a student admits to use within a certain period (for example, the past month) but denies any use within a broader period (for example, the past year). To gather evidence about probable causes, we examined the different patterns and checked for prevalence changes across time. To confirm the accuracy of substance-use responses, we created an indicator of a student's propensity to use based on several other questions. For example, the cigarette "propensity index" (PI) included the following questions among others: "Does your best friend smoke?" "Do you believe smoking relaxes you?" and "Do you enjoy doing something wrong just for fun?" Because these questions concern attitudes and others' usage, students have less incentive to hide the truth. For those with complete and consistent responses, the PI discriminated well across the different recency-of-use values (R^2=0.5)—an indication that the students' responses to the usage questions accurately reflected their propensity to use. We then used the PI to measure whether we could extract enough reliable information from the inconsistent students to estimate their usage accurately. After computing the average PI across students with similar response patterns, we compared the means of the inconsistent or missing groups with the means of relevant consistent groups.

LONGITUDINAL INCONSISTENCIES

Classification of Longitudinal Inconsistencies

Longitudinal inconsistencies differ from each other in several important ways: the nature of the inconsistency (type), the amount of use contradicted (severity), and the wave at which the inconsistencies occurred. Each factor needs consideration when we interpret longitudinal inconsistencies.

Type. We examined three types of longitudinal inconsistencies that we classified as "retractions," "modifications," and "new admissions."

- **Retractions (type 1):** A retraction occurred at a particular wave (2, 3, or 4) when a student denied any lifetime use after having admitted to use at an earlier wave.
- **Modifications (type 2):** A modification occurred when a student denied use in the past year following a wave in which he had admitted to use in at least the past month (three to nine months earlier).
- **New Admissions (type 3):** A new admission occurred when a student admitted use *for the first time* at wave 2, 3, or 4, but denied any use in the past year at that wave.

Severity. We defined an inconsistency's severity to distinguish cases that were explainable by relatively minor recall errors from those that suggested major rewriting of the facts. Thus, two people who admitted to use but then subsequently denied it would be viewed differently if one had previously smoked less than one cigarette and the other had smoked a pack a day. We determined the severity by examining additional information provided in related questions concerning the frequency and dosage of use.

Wave. For certain types of inconsistencies, the wave at which the inconsistency occurs affects the likelihood of recall errors. For example, because the time between waves 2 and 3 (nine months) is three times as long as for the other two intervals and covers the transition from seventh to eighth grade, we were more lenient toward time-dependent discrepancies (modifications, new admissions) occurring at wave 3.

Underlying Assumption

Our underlying assumption is that students who give at least one positive response have used the substance. To reach this decision, we studied responses to the age-at-first-use questions that appeared near the end of the first two questionnaires. Though a student might inadvertently *circle* an unintended response, he should be less likely to accidentally *write* an age. Thus, we used the written age at first use to validate the earlier circled responses.

Table 3.1

CIGARETTE AGE-AT-FIRST-USE RESPONSE FOR STUDENTS WHO ADMITTED USE AT WAVE 1, BY WAVE 2 BATTERY RESPONSE

Age at First Use (Wave 1 in yrs)	Deny Use at Wave 2	Admit Use at Wave 2
never	12.1	8.4
1-5	16.6	6.2
6-9	26.1	27.3
10	17.8	21.1
11	15.9	20.4
12	9.6	14.3
13	1.9	2.0
14	0.0	0.2
Overall	100.0[a]	100.0[a]

[a]Due to rounding, percentages may not add to 100.

As Table 3.1 shows, 88 percent of those who later denied lifetime use wrote in an age of first use. Moreover, the distribution of responses for students who admitted to use at wave 1 and then denied use at the following wave is largely similar to the distribution for those who consistently admitted to use at the first two waves. Because it would be difficult to distinguish between these groups based on their wave 1 responses, we can assume their responses are based on similar experiences. We thus have additional reason to believe the admissions of the inconsistent students. The largest discrepancy (16.6 versus 6.2) occurs for inconsistent students who said they had first tried cigarettes at a very young age, before they were six. If accurate, this use was probably limited to a puff or two — an incident that the student might not interpret as use when, and if, it is recalled at another wave.

The findings are similar when the use is denied at wave 1 rather than at wave 2 (see Table 3.2), and when it is for alcohol and marijuana (not shown). Because most inconsistent students at waves 1 and 2 provide positive and plausible age-at-first-use responses, we have further evidence to support the assumption that an admission of use, even if inconsistent, implies that a student has used the substance.

Table 3.2

CIGARETTE AGE-AT-FIRST-USE RESPONSE FOR STUDENTS WHO ADMITTED LIFETIME USE AT WAVE 2, BY WAVE 1 BATTERY RESPONSE

Age at First Use (Wave 2 in yrs)	Deny Use at Wave 1	Admit Use at Wave 1
never	11.9	8.4
1-5	12.6	6.4
6-9	30.5	28.2
10	20.5	18.4
11	15.9	19.4
12	8.0	16.2
13	0.7	2.9
14	0.0	0.3
Overall	100.0[a]	100.0[a]

[a]Due to rounding, percentages may not add to 100.

Retractions (Type 1)

The reasons for retracting an admission range from forgetting an initial taste of wine several years ago to trying to conceal heavy use. To identify the inconsistencies' probable causes, we examined the frequency and dosage of previously admitted use. Using the additional information, we further classified retractions into minor, questionable, and major inconsistencies:

- Minor—denied after having previously admitted light use (less than three times in one's lifetime or in the past year).
- Questionable—denied after having previously admitted moderate use (three to ten times in the past year or three to five days in the past month).
- Major—denied after having previously admitted frequent use (at least 11 times in the past year or six or more days in the past month).

Modifications (Type 2)

Modifications differ from retractions in that the student does not deny that he has ever used but instead modifies his statement about his most recent use. Because the student admits at least some past use, the type 2 inconsistency is more subtle than the type 1 and thus gives us less reason to question the reliability of the student who commits it. To distinguish among modification inconsistencies, we considered both the wave at which the discrepancy occurred and the frequency of use at the previous wave.

Students who denied use in the past year at wave 3 after admitting to recent use nine months earlier may either have been thinking in terms of the school year rather than the calendar year or had trouble distinguishing the past ten months from the past 12. Because of this extended time interval, we considered all wave 3 inconsistencies minor. For inconsistencies at waves 2 and 4, we assessed severity by examining the frequency of past month use at the previous wave. We thus divided the modification inconsistencies into three classes of severity:

- Minor—all inconsistencies occurring at wave 3, plus errors at wave 2 or wave 4 in which the student had used the substance only one or two days in the past month at the previous wave.
- Questionable—errors in recall at wave 2 or wave 4 in which the student had used the substance three to 19 days in the past month at the previous wave.
- Major—errors in recall at wave 2 or wave 4 in which the student had used the substance 20 or more days in the past month at the previous wave.

New Admissions (Type 3)

Given that waves have less than one year between them, students who change their use status from "never" to "ever" without acknowledging use within the past year are inconsistent. New-admission inconsistencies can occur for two very different reasons: 1) a student actually used the substance before the previous wave but failed to mention the usage at that wave; or 2) the student first used the substance between the two waves in question but incorrectly identified the recency of that usage. Analysis of the type 1 and 2 inconsistencies indicates that both these errors occur. Admission of use at a young age followed by denial causes retractions; if the pattern is reversed, which seems equally likely, the result is a new-admission inconsistency. Similarly, if a student initiates use soon after a data collection wave, the same sort of recollection error that would cause a modification could cause a new

admission. Distinguishing between these two categories is important because the latter deals with students who initiated use after the program's start.

Analysis of the data suggests that the predominant cause of the new-admission inconsistencies differs by waves. We used the supplemental data about age at first use to provide additional information for students who had denied use at wave 1 and admitted lifetime use three months later at wave 2. By comparing the age at first use to age at wave 1, we determined that 92 percent of these students had initiated use before baseline.[1] Other studies have noted that, particularly for illicit drugs, retrospective reporting may be more valid than concurrent self-reports if the subject fears legal ramifications (Collins et al., 1985). Thus, some students may have been hesitant to admit to use at baseline; however, once they felt confident no negative consequences would ensue, they admitted the earlier usage at the subsequent wave. We concluded that most of these students had probably used the substance before baseline and classified these discrepancies under the first category.

In contrast, we classified inconsistencies that occurred at wave 3 or 4 under the second category (initial use occurring between the two waves in question) because we considered it unlikely that students had used the substance more than a year ago, denied the use *at multiple waves*, and then finally acknowledged the use. In these cases, we believe it more likely that the inconsistent students had started using since the previous wave and incorrectly identified the recency. The negative responses to the age-at-first-use questions at the first two waves provide further evidence to support this decision.

In summary, we defined two categories of new admissions based on the wave at which the student first admitted lifetime use:

- Wave 2—student actually used the substance before wave 1 but denied the usage until wave 2.
- Wave 3 or 4—student first used the substance after the previous wave but incorrectly recalled the recency of that usage.

[1]The remaining 8 percent of new admissions at wave 2 had initiated use during the year they completed wave 1. Because we did not have precise date information, we could not determine whether they initiated use before or after the baseline survey.

IMPACT OF RESOLUTIONS ON TREATMENT EFFECTS

Because the study's primary goal is to analyze treatment effects, identifying whether inconsistencies and our resolutions affect the conclusions is essential. Problems that occur at random across experimental groups should not systematically influence any treatment effect results. A student who misreads a question, or commits some other inadvertent error, is as likely to be in one group as another. Thus, whatever biases arise in prevalence estimates should tend to cancel when experimental groups are contrasted to estimate treatment effects. In particular, tests for statistically significant differences will remain valid.[2] However, inconsistencies associated with other causes may have a greater effect. For example, assume that students in the treatment group retract previous use more than those in the control group because the program has made use less socially acceptable. Depending on how the inconsistent cases are handled, an analysis of the percentage of users by experimental group might exaggerate the program's positive effect on reducing drug use.

Because some inconsistent responses may indicate an unwillingness to admit one's usage, we looked for systematic differences between the experimental groups that might bias comparisons among them. To compare the groups, we performed two sets of tests. First, we used an analysis of variance (ANOVA) to compare the three groups (two treatments and one control) simultaneously. For the second comparison, we combined the two treatment groups into one group. We then performed t-tests to identify any differences between the control group and the combined treatment groups.

DEMOGRAPHIC DIFFERENCES

We also investigated whether the prevalence of inconsistent responses varied by demographic characteristics. Though we cannot conclude that groups with a greater percentage of inconsistencies are more likely to misreport intentionally, we may learn that certain traits are associated with a greater likelihood for inaccurate data. Because inconsistencies are more likely to occur among students who have used, we controlled for usage by stratifying the students by their baseline-use status—nonuser, experimenter, user (for definitions of the groups, see Ellickson et al. 1988). We made comparisons using chi-square tests between the different categories of gender, race/ethnicity, and scholastic performance.

---

[2]If inconsistencies of other problems reduce estimated prevalence proportionally in all experimental groups, the result would be reduced power to detect true differences. However, there would be no impact under the null hypothesis of no true difference among experimental groups.

IV. RESULTS

CROSS-SECTIONAL PROBLEMS

Incomplete Responses

Table 4.1 shows the prevalence of missing data within the battery of usage questions for each target substance. At waves 1 and 2, students should have skipped out of the item sequence only if they had never used. However, some students apparently misinterpreted the instructions and unconditionally omitted the remaining questions. After the skip pattern was removed, the most common reason for incomplete data occurred when a student denied lifetime use and felt no need to answer the questions that focus on recency of use. Thus, at waves 3 and 4, there is a greater percentage of missing data for marijuana, the substance with the lowest prevalence of lifetime use, than for cigarettes or alcohol.

Lack of time was a very minor cause of missing data for the substance-use questions. We designed the data collection sessions so that the first 15 to 20 minutes were spent on administrative work, saliva collection, and instructions—leaving 30 to 35 minutes for the students to fill out the 25 page questionnaire. At baseline, more than 90 percent of the participating students successfully completed the survey, while another 4 percent completed all the questions relating to substance use but failed to answer some demographic questions on the last few pages. We had designed the questionnaire so that almost all questions about use of the target substances—everything but age at first use—were in the survey's first section. Thus, all but five students reached the end of all three substance-use batteries. We used the age questions primarily to validate the earlier substance-use responses and thus purposely placed them farther back in the form.

Table 4.1

PERCENTAGE OF CASES WITH MISSING DATA, BY SUBSTANCE AND WAVE

Wave	Cigarettes	Alcohol	Marijuana
1	2.4	3.2	2.0
2	1.7	3.6	2.8
3	1.5	1.2	1.9
4	1.5	2.0	2.7

Of the students who reached the end of the survey at wave 1, 37 percent left at least one question blank that should have been answered. An examination of the relative frequency of blanks for different types of questions provides additional insight into the reasons for omitted responses. Almost all questions drew at least a background level of nonresponse equal to about one-quarter to one-half of that for the most frequently omitted questions. Those with the highest prevalence of nonresponse fall into two categories: substance-use questions and speculative questions. Between 1 and 3 percent of the students did not respond to items pertaining to the respondent's actual use. Some of the nonresponse may have been due to confusion with the skip patterns, since the percentage of missing substance-use data dropped considerably at the latter two waves. Speculative questions such as "About what percent of the eighth graders in your school do you think smoked one or more cigarettes in the last month?" were typically left blank more frequently than the substance-use items (1 to 5 percent). These patterns suggest that some students may have been hesitant about providing "self-incriminating" responses, but that nonresponse was more frequently attributable to uncertainty about the "correct" answer and other reasons.

Table 4.2 presents examples of incomplete responses to the battery of substance-use questions. At waves 1 and 2, the incomplete pattern in row 8 was most prevalent—probably caused by skip-pattern confusion—while row 10 represents the most common missing-data pattern at the subsequent waves. To determine the probable causes of missing data we can consider the strategies that would be used to complete the different patterns. Patterns in the table's upper section (rows 1 to 7) could be filled in by assuming that the data provided are accurate. In these cases, no obvious incentive for omitting the missing information exists, so we assume that the incompleteness does not reflect an intentional decision to withhold the truth. However, the patterns in the lower section (rows 8 to 11) present a problem because the missing response(s) occur in pivotal locations and thus could directly relate to use. For example, in row 10, the student reports that he has not used in the three most recent periods but does not respond to the lifetime-use question. If he had actually used, he might have chosen the safest alternative and left the question blank rather than lie or admit to use and face possible retributions. Depending on the wave and substance, between one-quarter and three-quarters of the cases with missing responses fall into this questionable category. Fortunately, this is at most 1.7 percent of the sample at any given wave.

Inconsistent Responses

Table 4.3 shows that for each substance, the percentage of inconsistent cases declined after the first two waves. This reduction seems attributable to two factors. Because of the skip pattern at waves 1 and 2, students who denied lifetime use and accurately skipped the following questions did not have the opportunity to be inconsistent within the battery. However, they still had to respond to the age-at-first-use question at the end of the survey. Discrepancies between those two questions accounted for nearly half of the inconsistencies at the earlier waves. With the exclusion of the age-at-first-use questions at the subsequent waves, all substance-use questions were located in the same section, reducing the potential for inconsistencies. Also, by wave 3, we can assume that the students had become more familiar with the questionnaire; they had also acquired greater maturity and comprehension skills over the intervening year. Though the percentages for cigarettes and alcohol are similar, those for marijuana are generally smaller. Because nonusers are less likely to give inconsistent responses and the prevalence of marijuana use is much lower, this result is not surprising.

Table 4.2

MISSING DATA RESPONSE PATTERNS

	Admitted Use			
	Life-time	Past Year	Past Month	Past Week
1.	Yes	No	—	—
2.	Yes	No	No	—
3.	Yes	Yes	No	—
4.	No	No	No	—
5.	—	Yes	No	No
6.	Yes	—	Yes	No
7.	Yes	Yes	—	Yes
8.	Yes	—	—	—
9.	Yes	Yes	—	No
10.	—	No	No	No
11.	—	—	—	—

Note: — denotes missing data.

Table 4.3

PERCENTAGE OF CASES WITH INCONSISTENT DATA, BY SUBSTANCE AND WAVE

Wave	Cigarettes	Alcohol	Marijuana
1	3.5	3.2	1.4
2	3.1	3.6	1.6
3	2.0	2.3	0.7
4	1.6	2.9	0.8

Table 4.4

INCONSISTENT RESPONSE PATTERNS

	Admitted Use			
	Life-time	Past Year	Past Month	Past Week
1.	No	No	No	Yes
2.	No	No	Yes	No
3.	Yes	No	No	Yes
4.	No	Yes	Yes	Yes
5.	No	Yes	Yes	No
6.	Yes	No	Yes	Yes
7.	Yes	No	Yes	No
8.	No	Yes	No	No
9.	Yes	Yes	No	Yes
10.	No	Yes	No	Yes
11.	No	No	Yes	Yes

Table 4.4 presents the possible inconsistent patterns that could occur across the four questions posed for each target substance. No one inconsistent pattern accounted for more than 1 percent of all battery responses, though those in rows 8 and 9 were most prevalent. In some cases only one response disagrees with the other three (rows 1 to 6). These inconsistent responses could have easily been caused by a small misunderstanding or an isolated accidental response. This type of minor discrepancy did not reduce our confidence in the ability to extract accurate information from those students.

The remaining patterns in the table have no obvious resolution. Those in rows 7 to 9 have two responses in conflict, either of which could be in error. The final two patterns have multiple inconsistencies that seem to represent randomness more than logic. Whether deliberate or careless, they cast serious doubt on the respondent's accuracy. As with the incomplete data, approximately half of the inconsistent cases (or less than 2 percent of all cases) involved such discrepant responses that we could not construct an accurate recency-of-use value.

The propensity index supports the belief that we can obtain reliable information from the minor inconsistencies (rows 1 to 6). If we assign a recency-of-use value based on the most obvious resolution (that is, changing the one discrepant value), we see that in general, the average PI for the resolved groups is most similar to the average PI for the consistent group with the same recency value (Table 4.5). This relationship holds across all waves.

LONGITUDINAL INCONSISTENCIES

Retractions (Type 1)

As Table 4.6 shows, retractions occur less frequently for marijuana use than for the other substances. This difference results from the relatively smaller percentage of students who had ever used marijuana (22 percent at wave 1 versus 55 percent for cigarettes and 77 percent for alcohol). If we omit students who reported no previous use (that is, they had

Table 4.5

AVERAGE PROPENSITY INDEX SCORES FOR CONSISTENT AND RESOLVABLE INCONSISTENT CASES, BY RECENCY OF USE

	Cigarettes		Alcohol		Marijuana	
Most Recent Use	Consist	Inconsist[a]	Consist	Inconsist[a]	Consist	Inconsist[a]
Never	20	25	44	47	7	7
Lifetime	34	40	50	50	23	20
Past year	54	58	58	53	44	18
Past month	70	71	74	67	63	68
Past week	88	107	80	97	78	67
Past two days	98	77				

[a]Includes only inconsistent cases with uniquely resolvable patterns.

nothing to retract), we see that the proportion of marijuana use retractions actually surpasses that for cigarette and alcohol use (Table 4.7). These percentages parallel those reported by Single, Kandel, and Johnson (1975): In their two-wave study, approximately 10 percent of the acknowledged users on the first survey denied use of the same substance at the second.

Table 4.8 combines inconsistencies across the waves and reclassifies them based on the recency and frequency of admitted prior use as defined in Sec. III. For each substance, about 80 percent of the retractions involved denials by students who had previously used once or twice at most, typically at least a year before baseline (rows 1a to 1c). Some students wrote in that this "trial" use consisted of sips or puffs. Changes about limited "experimenting" could very reasonably represent redefinitions of what constitutes use or even failure to remember it. For marijuana, forgetting experimental use seems less likely; redefining it seems more likely. Regardless of the cause, however, the implication of nonuse since previous admission seems accurate for most of the cases.

Table 4.6

PERCENTAGE OF STUDENTS WITH RETRACTIONS, BY SUBSTANCE AND WAVE OF THE RETRACTION

Wave	Cigarettes	Alcohol	Marijuana
2	3.8	6.0	2.1
3	5.2	7.1	2.5
4	7.5	8.3	4.2

Table 4.7

PERCENTAGE OF USERS AT PREVIOUS WAVE WITH RETRACTIONS, BY SUBSTANCE AND WAVE OF THE RETRACTION

Wave at Which Retraction Occurred	Cigarettes	Alcohol	Marijuana
2	8.5	8.5	13.4
3	8.1	7.0	9.2
4	6.5	5.3	8.5

Table 4.8

CLASSIFICATION OF RETRACTIONS

Description of Use Later Denied	Percentage of All Retractions		
	Cigarettes	Alcohol	Marijuana
1a. Light use a year or more before baseline	50	37	37
1b. Light use within year before baseline	10	18	15
1c. Light use after baseline[a]	23	28	25
2. Moderate use	12	14	7
3. Frequent use	6	4	7
Total	100	100	100

[a]In some cases, we cannot identify whether a student used only before baseline or before baseline and again shortly after. We have included these cases in category 1c. Approximately one-third of the cigarette and alcohol cases in 1c could belong in 1b, while 12 percent of these marijuana cases are indiscernible and may belong in 1b.

That the great majority of retractions are minor is also consistent with the findings of Single, Kandel, and Johnson (1975), who concluded that most of the students who later denied use had used the drug only once or twice in the past. We classified the responses of less than 5 percent of all students as moderate (row 2) or severe (row 3). The former category undoubtedly included students who truly felt that they had not had any noteworthy use plus those who intentionally chose to conceal their previous use. The latter group, which constitutes roughly 1 percent of all students, is probably predominantly composed of intentional concealers. These students were omitted from the analysis.

Modifications (Type 2)

The pattern for modifications mimics that for retractions. At each wave, proportionately fewer modifications occur for marijuana than for the other substances (Table 4.9). But the disparity disappears when we tabulate them as a percentage of students who admitted to use within the past month at the previous wave (Table 4.10).

Tables 4.9 and 4.10 also show that modifications at wave 3 are almost double the proportions at the other waves. These differences undoubtedly reflect the longer period before wave 3 (nine months) than before waves 2 and 4 (three months). Probably half of

these students at wave 3 are inconsistent because of a time discrepancy. While admitting prior use, they failed to place it within the last 12 months. Because the interval between the other two sets of waves is much shorter, memory lapse is less likely the cause for discrepancies at waves 2 and 4.[1]

As Table 4.11 shows, an average of 75 percent (61 - 92 percent) of students who denied use in the past year previously used on only one or two days. In contrast, 50 percent of all students who admit to use in the past month respond that they have used on at least three days. From these data, we infer that most inconsistent students inadvertently failed to recall the recency of the infrequent use rather than consciously denied it.

Table 4.9

PERCENTAGE OF STUDENTS WITH MODIFICATIONS, BY SUBSTANCE AND WAVE OF THE MODIFICATION

Wave	Cigarettes	Alcohol	Marijuana
2	0.8	0.7	0.3
3	1.3	1.8	0.6
4	0.7	1.1	0.6

Table 4.10

PERCENTAGE OF MODIFICATIONS AMONG STUDENTS WHO ADMITTED USE WITHIN THE PAST MONTH AT THE PREVIOUS WAVE, BY SUBSTANCE AND WAVE OF THE MODIFICATION

Wave	Cigarettes	Alcohol	Marijuana
2	5.7	3.5	7.1
3	8.7	5.6	10.4
4	3.4	2.9	5.2

[1]An alternative interpretation rests on the possibility that response errors may be more highly correlated when repeated measures are taken closer together and thus may account for the smaller percentage of inconsistencies at waves 2 and 4. If so, the data at wave 3 would provide the best reliability measure.

Table 4.11

COMPARISON OF NUMBER OF DAYS OF USE AT THE PREVIOUS WAVE AMONG STUDENTS WITH MODIFICATION INCONSISTENCIES

	Cigarettes			Alcohol			Marijuana		
Wave	1-2 (%)	3-19 (%)	20+ (%)	1-2 (%)	3-19 (%)	20+ (%)	1-2 (%)	3-19 (%)	20+ (%)
2	66	28	6	82	12	6	92	8	0
3	72	17	11	84	14	2	70	26	4
4	61	33	6	85	11	4	70	23	7

Studying discrepancies between responses about use in the past month and use in the past year, O'Malley, Bachman, and Johnston (1983) concluded that a major part of the discrepancy resulted from systematic underestimation of use in the past year. This work also supports our conclusion that most modification inconsistencies are caused by poor recall of limited use over the past year and should not be viewed as invalidating the data.

New Admissions (Type 3)

Table 4.12 presents the percentage of students at each wave with new-admission inconsistencies, showing once again that the numbers are very similar for cigarette and alcohol use (1.2 - 3.6 percent) and slightly smaller for marijuana use (1.0 - 2.3 percent). These percentages are approximately half the size of those found by Single, Kandel, and Johnson (1975). If we look at the new admissions in terms of a percentage of students who denied use at all previous waves (Table 4.13), we see an even greater disparity across substances. As Collins et al. (1985) also found, this type of inconsistency constitutes a

Table 4.12

PERCENTAGE OF STUDENTS WITH NEW ADMISSIONS, BY SUBSTANCE AND WAVE OF THE NEW ADMISSION

Wave	Cigarettes	Alcohol	Marijuana
2	3.4	2.4	1.3
3	3.6	3.1	2.3
4	1.5	1.2	1.0

smaller percentage of the nonusers of illicit drugs than of the nonusers of legal drugs. As with modifications (the other time-dependent inconsistency), up to twice as many discrepancies appear at wave 3 because of the extended time interval.

IMPACT OF ERRORS ON TREATMENT EFFECTS

Table 4.14 collapses the data across waves and presents the percentage of cases with inconsistencies compared across treatment groups. The major severity rows show the percentage of most-severe discrepancies, regardless of type, as defined in Sec. III. The two columns on the right indicate statistically significant differences. Though some significant differences exist, no pattern conveys strong evidence of a treatment effect. In particular, no meaningful pattern appears in the category in which we might expect to find a treatment bias (major severities). These findings agree with the conclusions of Collins et al. (1985), who found no significant differences between treatment and control groups when comparing inconsistencies between reported and recalled use for each of the three substances.

DEMOGRAPHIC DIFFERENCES

Tables 4.15 - 4.17 compare retraction inconsistencies by gender, race/ethnicity, and self-reported school grades. For all substances, and at all levels of baseline use, males were more likely to deny ever using after having previously admitted usage (Table 4.15). This was most apparent for cigarettes, where a statistically significant difference existed for all levels of use. Differences in use rates between males and females do not explain the phenomenon. Either females are more likely to tell the truth, do a more consistent job of hiding it, or are less likely to make careless errors. However, our data do not allow us to determine which explanation holds true.

Table 4.13

PERCENTAGE OF STUDENTS WHO DENIED AT ALL PREVIOUS WAVES WITH NEW ADMISSIONS, BY SUBSTANCE AND WAVE OF THE NEW ADMISSION

Wave	Cigarettes	Alcohol	Marijuana
2	6.4	9.3	1.6
3	7.9	16.0	2.9
4	4.2	9.3	1.4

Table 4.14

PERCENTAGE OF INCONSISTENCIES, AT ALL WAVES COMBINED, BY TREATMENT GROUP

Cigarettes	Treatment Group TLC	HEC	CON	Comparison of TLC vs. HEC vs. CON	Comparison of TX vs. CON
Retractions	11.5	12.2	9.9	—	[a]
Modifications	2.8	2.7	2.9	—	—
New admissions	8.1	8.6	8.9	—	—
Major severity	0.58	0.70	0.61	—	—
Alcohol					
Retractions	14.4	16.3	14.7	—	—
Modifications	3.2	3.7	4.2	—	—
New admissions	7.2	6.6	6.2	—	—
Major severity	0.64	0.76	0.38	—	—
Marijuana					
Retractions	5.3	7.7	5.1	[b]	[a]
Modifications	1.6	1.6	1.1	—	—
New admissions	4.6	4.3	4.8	—	—
Major severity	0.47	0.44	0.45	—	—

Note: TLC = Teen Leader Condition; HEC = Health Educator Condition; CON = Control Group; TX = TLC and HEC combined.

[a] $p < 0.10$.

[b] $p < 0.05$.

— no significant difference.

The patterns are somewhat less consistent for the other comparisons. For all substances, whites commit fewer retractions than do minorities (a combination of blacks, Hispanics, and Indians [Table 4.16]). These findings agree with those of Mensch and Kandel (1988) who found retraction twice as common among blacks and Hispanics as among whites. Though Asians are least likely to retract marijuana use, those who have used or experimented with cigarettes or alcohol at baseline are most likely to retract their use at subsequent waves. One possible explanation could be that nonwhites have a poorer comprehension of the English language. Comparing the inconsistencies by school grades, we see that the percentage increases as grades decrease among the baseline experimenters

Table 4.15

PERCENTAGE OF STUDENTS WITH RETRACTION INCONSISTENCIES, AT ALL WAVES COMBINED, BY GENDER AND BASELINE-USE STATUS

Baseline Use	Gender	Cigarettes	Alcohol	Marijuana
Nonusers	Male	5[a]	13[a]	1.4
	Female	2	9	0.9
Experimenters	Male	28[a]	25[a]	5[b]
	Female	19	20	4
Users	Male	11[a]	8	26
	Female	5	5	22

[a] $p < 0.05$.
[b] $p < 0.10$.

Table 4.16

PERCENTAGE OF STUDENTS WITH RETRACTION INCONSISTENCIES, AT ALL WAVES COMBINED, BY RACE/ETHNICITY AND BASELINE-USE STATUS

Baseline Use	Race/ Ethnicity	Cigarettes	Alcohol	Marijuana
Nonusers	White	3	10[a]	1.1
	Asian	4	8	1.0
	Minority	5	15	1.2
Experimenters	White	20[a]	19[a]	4
	Asian	36	29	4
	Minority	31	29	6
Users	White	6[a]	5[a]	23
	Asian	14	14	18
	Minority	12	11	26

[a] $p < 0.05$.

Table 4.17

PERCENTAGE OF STUDENTS WITH RETRACTION INCONSISTENCIES, AT ALL WAVES COMBINED, BY ACADEMIC PERFORMANCE AND BASELINE-USE STATUS

Baseline Use	Grade	Cigarettes	Alcohol	Marijuana
Nonusers	A	3[a]	9	1.3[a]
	B	3	10	0.5
	C	6	17	2.5
	D/F	11	0	0.0
Experimenters	A	22	19[a]	3
	B	23	22	5
	C	24	26	4
	D/F	31	28	9
Users	A	6	5	19
	B	8	7	23
	C	7	6	23
	D/F	7	2	16

[a] $p < 0.05$.

(Table 4.17). The small sample size for the nonusers with grades D or F (12 and 18 for alcohol and marijuana, respectively) may contribute to the counterintuitive patterns in those groups.

We performed the same analyses for students with modification inconsistencies, but did not find consistent or significant differences across gender, race/ethnicity, or scholastic performance (tables not shown). Perhaps recall errors where students admit at least some past use are more subtle and thus less likely to be associated with sociodemographic characteristics.

V. CONCLUSIONS

Obtaining accurate drug-use information through self-reports is a formidable task, and the challenge is even greater when the subjects are adolescents. Consequently, we invested much time and many resources at both ends of the study to maximize the data's accuracy. The preparation of a confidential, nonthreatening questionnaire minimized interpretation and concealment problems, while subsequent analysis of the data evaluated inconsistencies and identified unreliable data. This analysis yielded a series of conclusions about the data quality.

Most students gave honest and reasonably accurate responses. On average, less than 5 percent of the students provided incomplete or inconsistent data within a questionnaire. These cross-sectional rates reflect those found in other studies of self-reported drug use by adolescents (Barnea, Rahav, and Teichman, 1987; Single, Kandel, and Johnson, 1975). Moreover, the rates declined after we eliminated skip patterns from drug-use item batteries, suggesting that the problems largely reflected confusion and carelessness, rather than deliberate misrepresentation.

Severe inconsistencies between questionnaires were very rare. Although more than 40 percent of the students committed at least one longitudinal inconsistency across the four waves, over 95 percent of the discrepancies were minor—involving inconsistent reporting of experimental use (most of which occurred long ago) or errors in placing previously reported use within the appropriate 12 month period. Severe inconsistencies that reflected reversals of frequent use reports characterized only 5 percent of the total number of discrepancies. Our findings thus support those of other major studies which have concluded that the majority of inconsistencies are committed by students who have used infrequently. (Collins et al., 1985; Mensch and Kandel, 1988; O'Malley, Bachman, and Johnston, 1983; Single, Kandel, and Johnson 1975).

Accuracy levels for adolescent reports of alcohol and marijuana use are similar to those for cigarette use. The percentages of retractions, modifications, and new admissions for cigarette and alcohol use were nearly identical; after adjusting for the lower prevalence of use, the marijuana-use percentages were also in the same range. Moreover, independent verification of the cigarette reports with physiological data indicated that very few students lied about recent tobacco use. Ninety-five percent of those whose saliva tests yielded cotinine levels indicating recent use also admitted to using tobacco. The independent

validation of the cigarette reports, combined with the similarity in rates of inconsistent reporting across substances, supports our conclusion that the vast majority of students provided honest and accurate responses about all three drugs. However, although the rates are similar for all three substances, the cause for inaccurate reporting may differ between the legal and illicit substances. Fear of reporting a socially disapproved behavior may contribute more to reporting errors for marijuana (which is illegal for both adults and teenagers) than for alcohol and cigarettes (which are illegal for teenagers, but not for adults). While other studies have reached similar conclusions, our data do not allow us to judge whether this assumption is accurate.

Reporting errors do not threaten the validity of the treatment effects analysis. Because treatment effects analyses focus on changes in use after baseline, problems in recalling light use before wave 1 have minimal impact on the analyses. Inconsistencies involving initiation of use or denial of frequent or recent use could be cause for concern if the percentages varied by experimental group. We found little evidence that such problems occurred, concluding that inaccurate responses are independent of group and do not bias treatment effects. However, we did find evidence that use in the past year was underreported, even by otherwise consistent students. Thus finding statistically significant treatment effects may be more difficult.

We conclude that the overwhelming majority of Project ALERT students attempted to accurately report their drug use. By identifying and examining the incomplete and inconsistent cases, we have been able to eliminate the few students with completely inaccurate data from our analyses. With this confidence in the data's quality, we can now proceed with presenting prevalence estimates and treatment effect results.

Appendix

SENSITIVITY OF THE PREVALENCE OF LONGITUDINAL INCONSISTENCIES TO THE SAMPLE USED IN THE ANALYSIS

Because the students who failed to complete all four waves of data collection were not a random sample of those present at baseline, we should not assume that the subsample we have analyzed is representative of the original sample. Table A.1 presents the percentage of cases with longitudinal inconsistencies, stratified by the combination of waves attended. Of the students present at the first two waves, those who did not participate at either subsequent wave had the highest percentage of inconsistencies at wave 2 (row 1), while those who attended both the latter waves were least likely to commit inconsistencies at waves 2 and 3 (row 4). Not only is there a difference in percentages between students with "perfect" attendance and students without, but a difference also exists in the degree of missing data. By comparing the first two rows, we see that students who missed wave 3 but then returned for wave 4 were less likely to be inconsistent at wave 2 than those who missed both waves.

Table A.1

PERCENTAGE OF CASES WITH LONGITUDINAL INCONSISTENCIES, BY NUMBER OF WAVES ATTENDED

Waves Attended	n	Cigarettes			Alcohol			Marijuana		
		W1	W2	W3	W1	W2	W3	W1	W2	W3
1,2	432	3.2	9.0	[a]	2.8	11.3	[a]	2.8	6.9	[a]
1,2,4	190	3.7	5.8	[a]	3.2	9.5	[a]	0.5	3.7	[a]
1,2,3	791	2.9	5.6	14.3	2.3	6.9	14.7	1.4	3.5	9.5
1,2,3,4	4605	3.3	4.6	10.1	2.4	6.8	11.9	1.3	2.4	5.4

Note: n = sample size; w1 = wave 1; w2 = wave 2; w3 = wave 3. The W1 columns identify new-admission inconsistencies between waves 1 and 2. They are recorded under wave 1 because the denial occurs at that wave. The W2 columns identify inconsistencies at wave 2 resulting from retractions or modifications of use admitted at wave 1. The W3 columns identify all inconsistencies between wave 3 and earlier waves.

[a] Not present at wave 3.

At wave 1, the percentage of cigarette- and alcohol-use inconsistencies does not seem to relate to attendance at the subsequent waves. However, for marijuana use, the percentage (2.8 percent) of inconsistencies for those at only the first two waves was more than double that for any other combination of waves.

Alternatively, we can examine whether students who commit inconsistencies are more likely to miss subsequent waves than students who do not. Table A.2 shows that students with inconsistencies between waves 1 and 2 were more likely to miss a subsequent wave than those who provided consistent responses at the first two waves (27 versus 23 percent, $p < 0.05$). The pattern is similar when we look at nonresponse at wave 4 relative to whether an inconsistency occurred at wave 3 (18 versus 14 percent, $p < 0.05$). These relationships suggest that students who commit longitudinal inconsistencies might come from a more transient environment—a population more likely to have used or experimented with the target substances and thus more susceptible to committing inconsistencies. Though these numbers reveal that our report of the percentage of inconsistent cases may not be representative of the entire sample of surveyed students, our assessments regarding the cause of most discrepancies and the data's overall accuracy are still valid.

Table A.2

PERCENTAGE OF CASES ATTENDING SUBSEQUENT WAVES, BY LONGITUDINAL INCONSISTENCY STATUS

Waves 3 and 4 Attended	Inconsistent between Waves 1 and 2	
	Yes	No
Yes	73	77
No	27	23
Total	100	100

Wave 4 Attended	Inconsistent at Wave 3	
	Yes	No
Yes	82	86
No	18	14
Total	100	100

REFERENCES

Barnea, Zipora, Giora Rahav, and Meir Teichman, "The Reliability and Consistency of Self-reports on Substance Use in a Longitudinal Study," *British Journal of Addiction*, Vol. 82, 1987, pp. 891-898.

Bauman, Karl, and Clyde Dent, "Influence of an Objective Measure on Self-Reports of Behavior," *Journal of Applied Psychology*, Vol. 67, 1982, pp. 623-628.

Collins, Linda M., et al., "Agreement Between Retrospective Accounts of Substance Use and Earlier Reported Substance Use," *Applied Psychological Measurement*, Vol. 9, No. 3, 1985, pp. 301-309.

Ellickson, Phyllis L., R. M. Bell, M. A. Thomas, A. Robyn, and G. Zellman, *Designing and Implementing Project ALERT: A Smoking and Drug Prevention Experiment*, The RAND Corporation, R-3754-CHF, 1988.

Gfroerer, Joseph, "Influence of Privacy on Self-Reported Drug Use by Youths," in *Self-Report Methods of Estimating Drug Use: Meeting Current Challenges to Validity*, NIDA Research Monograph 57, Rockville, Md., 1985.

Johnston, Lloyd D., and Patrick M. O'Malley, "Issues of Validity and Population Coverage in Student Surveys of Drug Use," in *Self-Report Methods of Estimating Drug Use: Meeting Current Challenges to Validity*, NIDA Research Monograph 57, Rockville, Md., 1985.

Mensch, Barbara S., and Denise B. Kandel, "Underreporting of Substance Use in a National Longitudinal Youth Cohort: Individual and Interviewer Effects," *Public Opinion Quarterly*, Vol. 52, 1988, pp. 100-124.

O'Malley, Patrick M., Jerald G. Bachman, and Lloyd D. Johnston, "Reliability and Consistency in Self-Reports of Drug Use," *Internal Journal of the Addictions*, Vol. 18, No. 6, 1983, pp. 805-824.

Pentz, M. A., et al., "A Multicommunity Trial for Primary Prevention of Adolescent Drug Abuse," *J. of American Medical Association*, Vol. 261, No. 22, 1989, pp. 3259-3266.

Single, Eric, Denise Kandel, and Bruce D. Johnson, "The Reliability and Validity of Drug Use Responses in a Large Scale Longitudinal Survey," *Journal of Drug Issues*, Vol. 5, 1975, pp. 426-443.