Texas School Survey
of Substance Use
2000
Methodology Report and Validity Analysis

For the Texas Commission on Alcohol and Drug Abuse

Prepared by

James A. Dyer, Ph.D.
Bodhini Jayasuriya, Ph.D.
and
Kavita Jaswal

Public Policy Research Institute

Texas A&M University
College Station, TX 77843-4476

Introduction

The Public Policy Research Institute (PPRI), in conjunction with the Texas Commission on Alcohol and Drug Abuse (TCADA), conducted the seventh statewide survey of drug and alcohol use among Texas elementary and secondary students in the Spring of 2000. Originally implemented in 1988 as a component of a larger survey assessing substance use among the state's general population, the school survey has since become an ongoing, independent project. District surveys are offered every year with a statewide survey conducted every two years. The 2000 effort provides follow-up data reflecting changes over the past eight years in grades four through twelve.

The *Texas School Survey* project has two primary objectives. First, it serves to inform state and local policy-makers about the extent and nature of the substance use problem in Texas schools. Second, the statewide survey provides a standard of comparison for districts conducting local assessments of drug and alcohol use.

The purpose of this document is to describe the methodology used to administer the 2000 *Texas School Survey of Substance Use*. Following a brief introduction to the survey instrument itself, attention is then focused on sample selection and survey administration procedures. Next, methods for data processing and quality control are described and the report concludes with a review of standard error estimates.

Survey Instrument

Two versions of the 2000 *Texas School Survey of Substance Use* were developed and administered. The first was a six-page questionnaire designed for students in grades seven through twelve. The second was a simplified three-page instrument created for students in grades four through six. The elementary survey differs from the secondary survey in that it has simplified language and some complex questions were omitted. Elementary students were asked about only four types of substances including tobacco (cigarettes, snuff, and chewing tobacco), alcohol (beer, wine, wine coolers, and liquor), inhalants, and marijuana. Secondary students were asked about the same substances, as well as a broader range of illicit drugs including powdered cocaine, crack, hallucinogens, uppers, downers, steroids, ecstasy, Rohypnol, and heroin. Other sets of questions on both the elementary and secondary instruments were designed to assess behavioral correlates of substance use and students' perceptions of support available to help them cope with substance-related problems.

The questionnaire was in a format that could be scanned optically, similar to that used for standardized testing. It was designed for anonymous self-administration by students with the aid of a staff member to pass out the survey, read a common set of instructions, monitor the class during survey administration, and collect the instruments after they are completed. The survey instruments are included in Appendix A.

Survey Content

The 2000 *Texas School Survey of Substance Use* content remained essentially the same as that of 1998. In 1998, items on the elementary and secondary questionnaires were revised from those of previous years. The latter modifications, particularly with regard to the secondary questionnaire, were implemented in order to increase accuracy of response and to reduce the length and

repetition of the questionnaires. Those revisions were made to ensure compatibility with previous survey data.

Survey Sample

The sample of students for the 2000 survey was designed to be a random sample of all public school students between the fourth and twelfth grades in the state. In order to make administration practical, students were selected using a multi-stage stratified sampling procedure. This involved sampling districts, schools within districts, and classrooms within districts. All students in a sampled classroom were asked to participate in the survey.

For the 2000 survey, the sample design was modified from that of the previous years. The modifications are expected to increase the precision of survey results by taking advantage of newly available software packages that easily provide estimates of standard errors of the estimators resulting from complex sample designs. These computations were not possible prior to the arrival of these packages.

Selection of Districts

The primary analytic cluster was the school district since the approval needed to administer the survey had to be obtained at that level. First, the districts were stratified according to how urban the counties were in which they were located. Stratification along a variable deemed to be highly correlated with the characteristics of interest is a means of increasing the precision of estimates in complex sampling designs. The strata were formed as follows. First, 28 border counties formed a separate border stratum. Of the remaining counties, the most urban stratum involved counties with metropolitan populations of 1,000,000 or more, the next stratum with those between 250,000 and 1,000,000, and the third stratum with those metropolitan areas with less than 250,000. The remainder of the state constituted a final major stratum.

The strata were further subdivided by relative size of the districts, so that each stratum had a combination of large and small districts. Due to their large size relative to other districts, a total of nine districts were sampled with a probability of one. This means that these districts are always selected as part of the sample. They formed two of the substrata. The strata are listed in Table 1.

TABLE 1. Distribution of Selected Districts by Urban Class Size

Stratum	Group
1-P1	Large Urban Counties- 55,000 < enrollment (probability one districts)
1-A	Large Urban Counties- 20,000 < enrollment < 55,000
1-B	Large Urban Counties- 10,000 < enrollment <20,000
1-C	Large Urban Counties- enrollment <10,000
2-P1	Medium Urban Counties- 20,000 < enrollment (probability one districts)
2-A	Medium Urban Counties- enrollment <20,000
3-A	Small Urban Counties- 10,000 < enrollment

3-B	Small Urban Counties- enrollment <10,000
4-A	Non-Urban Counties- 5,000 < enrollment
4-B	Non-Urban Counties- enrollment <5,000
5-A	Border (28 counties)

Districts were selected for the state sample from the 11 strata, above, in the following manner. The nine districts that formed strata 1-P1 and 2-P1 were included with probability one. From each of the remaining strata, excluding the border stratum, a simple random sample of districts was selected. The districts in the border stratum were also selected with probability one in that all districts constituting that stratum were invited to participate. The districts that were selected in this manner are listed in Table 2.

If a non-probability one district refused to participate, it was replaced with another district selected at random from within the stratum.

TABLE 2. State Sample by Strata

Original State Sample	**Actual State Sample**
Stratum 1-P1 (n=5)	**Stratum 1-P1 (n=3)**
Houston	Houston
Dallas	Fort Worth
Fort Worth	Northside
San Antonio	
Northside	
Stratum 1-A (n=10)	**Stratum 1-A (n=3)**
Arlington	Arlington
Garland	Plano
Plano	Spring Branch
Pasadena	
Alief	
Lewisville	
Klein	
Spring Branch	
Conroe	
Carrolton-Farmers Branch	
Stratum 1-C (n=6)	**Stratum 1-C (n=3)**
Grand Prairie	Grand Prairie
Judson	Judson
Harlandale	Denton
Edgewood	

Denton
Duncanville

Stratum 1-D (n=11)	**Stratum 1-D (n=5)**
La Porte	Cleburne
Coppell	Lake Worth
Cleburne	Royse City
Carroll	Little Elm
Midlothian	Ennis (Alt)
Lake Worth	
Rose City	
Little Elm	
Ft Sam Houston	
Melissa	
Ennis (Alt)	

Stratum 2-P1 (n=4)	**Stratum 2-P1 (n=1)**
Austin	Austin
Corpus Christi	
Round Rock	
Beaumont	

Stratum 2-A (n=9)	**Stratum 2-A (n=2)**
Pflugerville	Tuloso-Midway
Nederland	Georgetown(Alt)
Port Neches-Groves	
Flour Bluff	
Lumberton	
Tuloso-Midway	
Orangefield	
Lago Vista	
Georgetown (Alt)	

Stratum 3-A (n=5)	**Stratum 3-A (n=2)**
Lubbock	Ector County
Ector County	Midland
Midland	
Abeline	
Tyler	

Stratum 3-B (n=6)	**Stratum 3-B (n=3)**
Temple	Pine Tree
Longview	Burkburnett
Pine tree	Arp
Burkburnett	
Highland Park	
Arp	

Stratum 4-A (n=5)	**Stratum 4-A (n=3)**
Victoria	Victoria
Nacogdoches	Nacogdoches
Sherman	Sherman
Greenville	
Kingsville	

Stratum 4-B (n=12)	**Stratum 4-B (=6)**
Beeville	Beeville
Stephenville	Stephenville
Wharton	Bruceville-Eddy
Graham	Schliecker County
Daingerfield-Lone Star	Honey Grove
Rice Cons	Beckville
Bruceville-Eddy	
Tahoka	
Schleicher County	
Honey Grove	
Joaquin	
Beckville	

Stratum 5-A (All districts in foll. counties)		**Stratum 5-A (participants)**	
Brewster	Kinney	Brewster(1)	Webb(1)
Brooks	La Salle	Cameron(4)	Willacy(3)
Cameron	Maverick	Duval(2)	Zapata(1)
Culberson	Pecos	El Paso(5)	
Dimmit	Presidio	Hidalgo(12)	
Duval	Reeves	Hudspeth(2)	
Edwards	Starr	Jeff Davis(1)	
El Paso	Terrell	JimHogg(1)	
Hidalgo	Uvalde	Kinney(1)	
Hudspeth	Val Verde	Reeves(1)	
Jeff Davis	Webb	Starr(2)	
Jim Hogg	Willacy	Terrel(1)	
Jim Wells	Zapata	Uvalde(1)	
Kennedy	Zavala	Val Verde(2)	

Obtaining cooperation from those districts that were randomly selected for the state sample when the selected district did not plan to do a local survey was sometimes a problem. Yet, it was critical to get data from as many of the originally selected districts as possible. Some state sample districts that were initially hesitant were persuaded to cooperate by the use of incentives. The various incentives used included waiving participation and sampling fees, offering to discount the

fees for participating the following year, paying all shipping costs, and discounting campus level analyses fees.

Seventy of the original 139 selected districts participated in the study. Sixty-nine districts were not able to participate, and most declined due to the lack of time and resources involved in survey administration. Many districts were preparing students for TAAS testing and expressed concerns about diverting resources away from that preparation. In lieu of the declining districts, an additional two districts were included in the final sample.

Out of the total of 70 districts that agreed to participate, the statewide sample consisted of 69 secondary and 69 elementary districts (See Table 2). One district, Fort Davis ISD, submitted only secondary data and one district, Tuluso-Midway ISD, provided only elementary data. The cooperation rate was 58 percent, with rates ranging from 22 to 60 percent. The cooperation rate was lowest for smaller districts in medium sized urban counties (Strata 2), however there were no consistent differences in cooperation rates between larger and smaller districts. In general, there was a trend for non-urban districts to have low cooperation rates (See Table 3). A total of 64 percent of the students in the original sample were in the final sampling frame.

TABLE 3. Cooperation Rate by Strata

	Strata 1A	**Strata 1C**	**Strata 1P1**	**Strata 1D**	**Strata 2A**	**Strata 2P1**	**Strata 3A**	**Strata 3B**	**Strata 4A**	**Strata 4B**	**Strata 5-A**
Total Cooperation Rate (58%)	30%	50%	60%	45%	22%	25%	40%	50%	60%	50%	60%

Participation of Border School Districts

In order to allow further analysis of substance use among students living on the Texas-Mexico border, school districts along the border were encouraged to participate in the 2000 *Texas School Survey*. The survey was offered free of charge to border districts, and data was collected from a broadly defined 28-county area. Subsequent analysis will focus on not only the larger border area, but also concentrate on a more strictly define 13 county border region.

The border sample was designed to collect data from approximately 315,000 students. Ninety school districts from 28 counties participated in the survey. A list of participating school districts is found in Appendix B. Fourteen of the border districts surveyed were included in the state survey sample.

Allocation of Surveys among Districts

The state survey sample was designed to collect data from a minimum sample of about 5,555 students per grade, however, many districts chose to survey more than the minimum number of students specified in the state sampling plan. Some extremely small districts received somewhat more than a strict proportional allocation because, while the data was technically only needed from one or two students per grade, the survey was administered to the entire classroom. Similarly, in a few extremely large (urban) districts, fewer students were need for accuracy than would result from a true proportional allocation. All surveys submitted from a cooperating district

were included in the sample. Accordingly, in the final analyses, the data were weighted to provide an accurate proportional allocation.

Thus, although we had estimated that the state sample would include approximately 50,000 students, it actually included 81,875 elementary students and 127,380 secondary students (See Table 4). This significantly improves the accuracy of estimates.

TABLE 4. Number of Surveys Included in State Sample

	Total Non-blank Surveys	Number of Useable	Number Rejected*	Percent Rejected
Secondary	136,014	127,380	8,634	6.3%
Elementary	82,827	81,875	952	1.15%
Total	218,841	209,255	9,586	4.38%

*Surveys were rejected because the responses indicated exaggeration or the survey could not be matched to a sampled school and grade.

Allocation of Surveys among Classrooms and Campuses

Once the number of surveys to be administered in each district was established, the next step was to determine the number of classrooms to be surveyed per grade. This was achieved by dividing the number of questionnaires per grade (ascertained for each district using proportional population calculations) by the average number of students per class---20 for grades four through six, 22 for grades seven through twelve. The result of this computation indicated the total number of classes to be surveyed. These classes were selected so that as many different campuses as possible were in the final sample. Ideally, the classrooms surveyed were evenly distributed across all campuses in the district. If there were more campuses containing a given grade than classrooms needed, then a simple random selection procedure was used to determine which campuses would be sampled. In general, once a campus was selected, all relevant grades at that campus were surveyed. Therefore, campus selection was not independent between grades.

TABLE 5. Survey Distribution by Grade

	Grade	Number of Usable Surveys	Percentage
Elementary	4^{th}	26,633	32.53%
	5^{th}	26,711	32.62%
	6^{th}	28,531	34.85%
		81,875	100%

Secondary	7th	25,272	19.84%
	8th	26,354	20.69%
	9th	22,715	17.83%
	10th	20,977	16.47%
	11th	15,931	12.51%
	12th	16,131	12.66%
		127,380	100%

Selection of Classrooms within Campuses

After the total number of classrooms to be surveyed in each grade at each campus was determined, it was necessary to identify specific classrooms. This selection procedure was performed by campus personnel based on a set of guidelines provided by PPRI (illustrated in Appendix C). Members of campus staff were asked to make a list by grade (according to teacher's last name or some other convenient method) of all classes held during a selected class period. They were then instructed to use a random number table to select the exact classes to survey in each grade.

Other Sampling Considerations

Some school districts sampled all students in all or some of the grades. In these districts, the methodology outlined above did not apply to the grades sampled at 100 percent. In Houston and Austin, the district used a list of all students from which to conduct a random sample of the students. Therefore, there are no campuses and classrooms sampled in these districts.

Survey Administration Procedures

Districts selected for inclusion in the state sample were notified about the project via letter and were sent a descriptive brochure, illustrated in Appendix D. State sample districts that planned to administer a local drug and alcohol survey had virtually no procedural changes resulting from their involvement in the statewide project. In those districts that surveyed grades four through twelve, sufficient data was collected from all relevant campuses to meet the data collection needs of the statewide survey. These districts benefited from their inclusion in the state survey project because they were not charged for the surveys that became part of the state database. The larger number of surveys from these districts were weighted down so that their contribution to the final sample was in correct proportion.

In those instances where state sample districts were collecting local data for an incomplete combination of grades, or where they were not conducting local surveys at all, the campus and classroom selection procedures described above were applied. Arrangements for giving the survey were established on an individual basis with these districts. Since those not doing local surveys did not stand to gain directly from having the survey administered in their district, an effort was made to be as accommodating as possible. PPRI was able to arrange survey administration in the selected schools and classes by school personnel.

In Houston and Austin, the district uses the computer to draw a random sample of all students. On each campus where the students are located, the students are requested to go to a specified room where the survey is conducted. Once in the room, the survey is conducted, as it would be in a classroom in the other districts.

Relevant personnel in the selected districts and campuses were provided with complete instructions and materials necessary to administer the survey (see Appendix E). Classrooms were selected randomly by PPRI based on information from a computer printout from the district or Campus Information Form. Teachers in selected classrooms were given a script to read so that all students would receive a standardized set of instructions. Teachers were also asked to complete a Classroom Identification Form that provided data on the number of students that should have taken the survey but were absent, and the number that were present but failed to complete the survey. This information was useful for computing error estimates. After the surveys were administered in each classroom, they were sealed in an envelope along with the Classroom Identification Form. The envelopes from all participating classrooms were collected and returned to PPRI.

Data Entry and Analyses

As noted earlier, the format of the survey instruments enabled them to be scanned optically. Upon receipt at PPRI, the instruments were logged in, coded, and scanned by staff or trained personnel.

Exaggerated Responses

Because the *Texas School Survey* data is based entirely upon respondents' description of their own behavior, it is inevitable that some students will under- or over-report their use of drugs or alcohol, and to the extent possible PPRI attempted to identify and eliminate data from those respondents. Two checks were incorporated into the data analysis program to identify exaggerators. First, both elementary and secondary students were asked about their use of a false drug call "cosma." Data from students claiming to have used this substance were considered suspect and dropped from the analyses.

Second, checks were run to identify any students claiming extremely high levels of drug and alcohol use. Unbelievable high substance use for elementary students was defined as the use of five or more substances, 11 or more times in the past year or over a lifetime. Secondary students were defined as exaggerators based on the following criteria: (1) students reported that they had five or more drinks of two or more beverages every day; (2) students reported that they had consumed three or more alcoholic beverages every day; or (3) students reported that they used four or more drugs (other than cigarettes, alcohol, or steroids) eleven or more times in the past month. As in those cases where students reported using "cosma," data from students reporting exaggerated use were also dropped from the analyses. Less than two percent (1.15%) of the total elementary sample exaggerated. The percentage of secondary school students who exaggerated (6.3%) was more than thrice that of elementary students.

Unreported Grade Levels

When students failed to report their grade level, it was impossible to determine unequivocally in which grade these students' data should be analyzed. When a grade level was missing, an estimate of the grade was made based on the students' age and the data were retained. Table 6 identifies the range of students' ages and the corresponding grade levels that were assigned. If both grade and age were missing, the data were dropped from the analyses.

TABLE 6. Age-Based Grade Assignments.

Age	Elementary Grade Level	Age	Secondary Grade Level
9	4th Grade	12	7th Grade
10	5th Grade	13	8th Grade
11	6th Grade	14	9th Grade
		15	10th Grade
		16	11th Grade
		17 or older	12th Grade

Quality Control Measures

To ensure the quality of the statewide survey data, a number of internal checks were put into place to guide survey processing. First, a quality control analyst oversaw the implementation of all pre- and post-analysis quality control procedures. As the following paragraphs describe, many aspects of PPRI's plan for quality control were embedded in automated procedures. However, there is no replacement for human oversight. The quality control analyst monitored and tracked the processing of each district's surveys from the initial mailing through the production of the final state report. Responsibilities included ensuring that surveys were properly coded and scanned and checking for anomalies in the final table of results.

In addition to the safeguards resulting from careful project oversight, there were also a number of procedural checks against error. For example, there was a possibility, however remote, that after the bindings of a set of survey instruments were cut, the instruments could be dropped or otherwise placed out of order. If this occurred, it is conceivable that some pages of data could have been read into the incorrect computer record. To resolve this problem, each instrument used in the 2000 survey was printed with a five-digit "litho-code" number. With this coding process, every page of a given instrument is printed with the same scannable number, but a unique number is assigned to every instrument. By using the litho-code, when each page of an instrument is scanned it will automatically be read into the correct computer record. In this way, even if the pages from different instruments were shuffled together and read randomly, all data derived from the same instrument would automatically be read to the same data record.

Litho-coding also enabled PPRI to confirm that data from every survey instrument read was associated with the correct district. Survey instruments were mailed to participating districts in consecutive order. By recording the beginning and ending instrument numbers going to each district, PPRI was able to check the litho-codes scanned for a given district. In this way, any stacks of data that could potentially have been inadvertently mislabeled could be easily identified.

Programming checks were also incorporated into the data analysis program by cross-analysis. That is, the same data was run in several different ways using existing programs, and program outputs were then compared for consistency. Confidence is high that these quality control features will ensure valid and reliable survey findings.

Weights, Standard Errors, and Confidence Intervals

Weights were applied to each case based on the strata (i.e., Urban Class I though IV), district, and campus. The weights were applied so that the aggregation of students in each campus, district, and strata reflected their proportions in the actual district, campus, and strata populations. The formulae use to determine these weights are presented in Appendix F.

Standard errors and confidence intervals were estimated for each grade and the aggregation. The formulae used are presented in Appendix G. The table of standard errors and confidence intervals for 30 day and lifetime use of substances by grades are presented in Appendix H.

Item Response Analysis

As with any survey, there were potential threats to the validity of the conclusions drawn from the data. Therefore it was important to examine the ways in which students' were responding to the questionnaire. Following the collection and TCADA approval of the data, all of the items on the survey were analyzed to assess the integrity of the data. We were specifically interested in exploring potential misinterpretation of questions, dishonest responses, and inattention to the survey questions and instructions.

Separate analyses were conducted for the total sample of elementary and secondary school survey responses. Additional analyses, exploring potential ethnic and grade-level differences were also conducted for the statewide secondary instrument.

Overall, the vast majority of students in both elementary and secondary schools appeared to have provided valid responses to the 2000 Texas Schools Survey of Substance Use. Few Students were classified as giving exaggerated responses. Likewise, any inconsistency that occurred was generally most likely due to inattention to survey instructions and questions, misinterpretation of the questions, or fatigue. Specific findings of the item analyses are highlighted below.

Elementary Survey

- Some students used the "Never heard of" and "Never used" response options interchangeably.
- Students who responded inconsistently about substance use were more likely to have initially reported no use and then acknowledged use on a later question, than to have cited use and recanted the use later in the survey.
- Questions at the end of the survey were somewhat more likely to be left unanswered than were those at the beginning.
- Students began answering most items that contained questions about multiple drugs, however, they routinely neglected to finish the item and answer questions about the final few drugs on the list.

Secondary Survey

- The largest percentages of inconsistent responses were most likely due to the survey's use of different terms for the same category substances across questions (i.e., cigarettes versus tobacco products, or spray paint versus inhalants).
- Other inconsistencies may be attributable to different interpretations of "use". Some students appear to interpret use in an answer as "regular use", whereas others seem to cite "use" when they may mean that they have "tried" a substance.
- Very few students who reported substance use in the past 30 days early in the survey subsequently denied use of the substance in later questions about the past 30 days.
- In contrast to the elementary students, secondary students (across all grades) were generally more likely to report use of a substance and later deny it, than visa versa.
- Asian and Caucasian students were more likely to respond consistently than students from other ethnic backgrounds.
- Students were more likely to leave questions at the end of the survey unanswered than those at the beginning.
- Some groups of questions were largely ignored by fairly large percentages of respondents.

Conclusion

The *Texas School Survey* has become a valuable policy tool for both state and local educators and policy-makers. The survey, performed every two years, provides timely and relevant information about current drug and alcohol use patterns among young people enrolled in the Texas' public schools. Furthermore, longitudinal analysis can provide insight into changes in drug and alcohol prevalence over time. As was noted in the introduction, every state survey culminates in a TCADA publication providing an overview of findings to date. Data is also available for independent analysis by policy-makers and academicians.