



**Washington State
Healthy Youth Survey**

Bias Analysis 2004

2004 Healthy Youth Survey Data

WASHINGTON STATE DEPARTMENT OF HEALTH

Healthy Youth Survey

Bias Analysis



Prepared by Washington State Department of Health

For more information contact:

Washington State Department of Health
TEL: 877-497-7111 (Toll Free)
Email: healthy.youth@doh.wa.gov

DOH Pub 160-180

For people with disabilities, this document is available on request in other formats. To submit a request, please call 1-800-525-0127 (TDD/TTY call 711).

Bias 2004

Overview

Survey responses are often used to estimate the frequency of behaviors or other characteristics in a population larger than those who actually complete the survey. Thus, while only a portion of Washington State public schools were sampled for the Healthy Youth Survey in 2004, we would like to use the student responses from these schools to characterize all 6th, 8th, 10th and 12th graders in Washington. This is only possible if those who participated in the Healthy Youth Survey are not different in their behaviors from those who did not participate. If they are different, we say that the survey is biased and we are then limited in our ability to generalize the results to all students. Bias represents systematic error and is different from the random fluctuation that is measured by confidence intervals.

From the analysis of bias presented below, we conclude that the results of the 2004 Healthy Youth Survey can be generalized to all public school students in 6th, 8th, 10th and 12th grades. However, caution should be exercised in using questions answered by 8th grade students that were asked at the end of the non-optional portion of the questionnaires. For the optional "tear-off" questions, there does not seem to be bias even though they were at the end of the questionnaire. While the reasons for this apparent discrepancy are unknown, completing the tear-off was decided at the school level, while failure to complete the survey was at the individual level.

Potential Sources of Bias in the Washington State Sample of the 2004 Healthy Youth Survey

The primary potential sources of bias in the 2004 Healthy Youth Survey include school participation bias, student response bias, "tear-off" bias, and survey completion bias:

- A random sample of schools, by grade, was asked to participate in the state sample of the Healthy Youth Survey. School participation bias would occur if there were differences between students in schools that participated and students in schools that chose not to participate.
- Among schools that did participate in the Healthy Youth Survey, student response bias would occur if there were differences between students that completed the survey and those that did not complete the survey.
- Not all participating schools completed the optional "tear-off" questions at the end of the survey. There may be bias if students in schools that participated in the "tear-off" were different than those who chose not to participate in the "tear-off".
- Bias might be present on the last items on the survey due to some students being unable to complete the questionnaire in the time allotted.

Methods of Assessing Bias

In order to assess possible bias, we compared

- characteristics of participating and non-participating schools, among those schools that were randomly selected for the state sample;
- weighted to non-weighted responses for Healthy Youth Survey questions (data weighted for student response and alternative school participation);
- the responses of students who completed the optional "tear-off" items to students who did not complete these items, based on other items that were completed by all students; and

- the responses of students who completed the last 30 items on the survey (not including the optional items) to students who did not complete these items, based on other items that were completed by all participating students. These analyses focused on Form B because survey non-completion was more pronounced for form B than the other forms.

To examine whether there were differences as described above, t-tests for continuous items and chi-squares tests for dichotomous items were conducted. The predictor variables were school participation (participated/refused), weighted/non-weighted Healthy Youth Survey student responses, student completion/non-completion of the optional "tear-off" questions, and student completion/non-completion of the questions at the end of the survey. Outcome variables depended on the type of bias being assessed and are discussed below. Analyses were generally conducted by grade and for all grades combined, although for some analyses, data were not available for all grades. Differences were considered statistically significant if the probability of finding a difference as large as the one measured would be expected to occur fewer than five times out of 100 (i.e. $p < 0.05$) by chance alone. Consideration of chance findings due to multiple comparisons were discussed, when needed, below. Additionally, where significant or marginally significant differences were found, the possible role of alternative school under-representation was examined.

Results and Conclusions

Characteristics of schools and students

Health Youth Survey School Participation and Student Response Rates

School participation rate ¹

Among schools that were randomly selected for the state sample, participation rates were 79.6 percent for 6th grade, 83.6 percent for 8th grade, 85.5 percent for 10th grade and 82.1 percent for 12th grade (Table 1). There was no significant difference in participation by grade ($p = 0.60$).

Grade	Schools Recruited	Participated	% Participated
6	103	82	79.6
8	67	56	83.6
10	69	59	85.5
12	67	55	82.1

Total Response rate ²

The total response rates before data cleaning were 69.1 percent for 6th grade, 76.2 percent for 8th grade, 61.7 percent for 10th grade and 51.5 percent for 12th grade. When only valid surveys were considered, the total response rates were 67.6 percent for 6th grade, 72.8 percent for 8th grade, 58.1 percent for 10th grade, and 48.3 percent for 12th grade (Table 2). Only valid surveys are used for the bias analysis.

The total response rates for valid surveys were significantly different by grade ($p < 0.001$). The total response rate was lower than the school participation rates due to student absences and refusals. In addition, 12th grade student response rates were affected by programs that involve students going to junior colleges or elsewhere off campus.

Table 2. Total Response Rate by Grade, HYS 2004 ²			
Grade	Enrolled in Recruited Schools	Completed Surevey (%)	Valid Survey (%)
6	11,623	8,029 (69.1)	7,862 (67.6)
8	11,627	8,858 (76.2)	8,466 (72.8)
10	13,879	8,557 (61.7)	8,059 (58.1)
12	12,157	6,260 (51.5)	5,876 (48.3)
Total	49,286	31,794 (64.5)	30,263 (61.4)

Bias Analysis

Due to the relatively high total response rate for grades 6 and 8, reasonable confidence exists that the results for these grades are generalizable to students in public schools statewide. The lower total response rates for grades 10 and 12 introduce the possibility of bias, due to possible differences between the students that did and did not complete the survey, or differences between participating and refusing schools.

Previous analysis of the participating and non-participating schools in the 2002 HYS found that alternative schools were less likely to participate in the survey compared with regular schools, and that differences in school characteristics were primarily due to poor representation of alternative schools.

These analyses compare participating and non-participating schools by alternative school status, school size, free and reduced lunch, minority enrollment, Washington Assessment of Student Learning (WASL) test scores, graduation rates, drop-out rates, and answers to selected questions from the Iowa Test of Basic Skills (ITBS) survey. All school data were obtained from the Washington State Office of Superintendent of Public Instruction ³ (OSPI).

Alternative School Participation

The school participation rates for grades 8, 12, and all grades combined (6, 8, 10 and 12) were significantly lower for alternative schools compared with regular schools (Table 3, data for individual grades not shown). These results indicate that alternative schools were under-represented among participating schools, especially for grades 8 and 12.

Table 3. Participation Rates for Alternative and Regular Schools, All Grades				
Type of school	Recruited	Participated	% Participated	p ~ value
Alternative school	30	19	63.3	
Regular school	208	172	82.7	
All Schools	238	191	80.3	p = 0.01

For grades 8, 10, 12, and all grades combined, total response rates were significantly lower for alternative schools compared with regular schools (Table 4, data for individual grades not shown). The total response rate for grade 6 alternative schools was significantly higher than the total response rate for grade 6 non-alternative schools (data not shown). These results indicate that alternative school students were under-represented in the survey participation for grades 8, 10, and 12.

Table 4. Total Response Rates for Alternative and Regular Schools, All Grades ⁵				
Type of school	Students Enrolled in Recruited Schools	Valid Survey	% Valid Survey	p ~ value
Alternative student	2,045	635	31.1	
Regular student	47,241	29,388	62.2	
All Students	49,286	30,023	60.9	p < 0.001

Small School Participation

The median number of students in grades 8 and 12 was significantly greater in the participating schools compared with the schools that did not participate ⁶. However, the differences did not remain after alternative schools were taken out of the analysis ($p > 0.19$, data not shown). It appears that the difference in school enrollment by participation was primarily due to alternative school status.

Other School Characteristics

Percent receiving free/reduced price lunches

The percent of children participating in the free or reduced-price lunch program provided an

estimate of socioeconomic status. There were no significant differences in the percent of students receiving free and reduced lunch by participating and non-participating schools (p 's > 0.38 for grades 6, 8, 10, 12, and combined grades).

Percent minority

There were no significant differences in the percent of minority students by participating and non-participating schools (p 's > 0.11 for grades 6, 8, 10, and 12).

10th grade WASL scores

There were no significant differences by participating and non-participating schools in the percentage of grade 10 students meeting the WASL standards in writing, reading, math, science, or meeting three of these standards. None of the tests approached significance (p 's > 0.12).

Graduation rates

There was no significant difference in the graduation rates of grades 10 and 12 students between participating and non-participating schools (p 's > 0.36).

Dropouts/status unknown

There was no significant difference between participating and non-participating schools in the percentage of students in grades 9-12 who dropped out or were of unknown status for the grade 10 sample ($p = 0.95$).

In grade 12, schools that did not participate had a significantly higher percentage of students who dropped out or were of unknown status, compared with those schools that did participate ($p = 0.04$). However, when alternative schools were taken out of the analysis, the difference was no longer significant ($p = 0.10$).

ITBS questions

A survey administered along with the Iowa Test of Basic Skills (ITBS) achievement tests in grade 6 provided additional information that was used to compare participating and non-participating schools:

- Students answering "yes" to the question "Do you have a computer in your home?" (yes/no);
- Students answering "yes" to the question "Have you attended any other school during this school year?" (yes/no);
- Students answering "B" or "C" in response to the following question: "How often is English spoken in your home?" (A. Only English is spoken in my home; B. Sometimes another language is spoken; C. Another language is spoken more often than English).
- Students answering "C" or "D" in response to the following question: "Do you feel safe at school?" (A. Always; B. Most of the time; C. Some of the time; D. Never).

None of the tests approached significance (p 's > 0.20).

Effect of Student Non-Response and Under-Representation of Alternative Schools

These analyses indicate that alternative schools were under-represented in the state sample and as a result, participating schools had significantly higher enrollment and a lower number of students who were drop outs or of unknown status. Not only were alternative schools more likely to refuse participation than regular schools, but also, among alternative schools that did participate, the proportion of enrolled students responding to the survey was lower than among regular schools (Table 4). This information suggests that the HYS results may provide under-estimates for some behaviors that differ between students in alternative and regular schools.

However, whether statewide rates were affected depends on the number of students in the alternative schools. Alternative schools tended to be smaller than regular schools. If the

number of students in alternative schools was small, even though students in alternative schools were under-represented, it might not have a major effect on statewide rates. Alternative school students in grades 10 and 12 comprise 5.5% of the total number of grade 10 and 12 students in the Healthy Youth sample.

Thus, in order to examine the possible effect of under-representation of alternative school students on statewide estimates, student-level HYS data were weighted both by student response and school participation rates and re-analyzed. These analyses were completed for grades 10 and 12 because most alternative schools were at the secondary level. Specifically, the weight for each student equaled the inverse of the student response rate for that school (calculated by dividing the number of valid surveys by the enrollment for that school) multiplied by the inverse of the participation rate for that type of school (alternative or regular). Thus, a regular school with low student response rate had a greater weight than a regular school with a high student response rate. An alternative school with an identical student response to a regular school had a greater weight because the weight was multiplied by the inverse of the school participation weight (which was greater for alternative schools).

The items that were compared included:

- Watch TV 2+ hrs/day
- Drink 2+ sodas/day
- Meet recommendations for vigorous activity
- Eat dinner with family most of time or always
- Overweight or at risk for overweight
- Smoked cigarettes in past 30 days
- Drank alcohol in past 30 days
- Used marijuana in past 30 days
- Gang membership in past year
- Seriously considered suicide in past year
- Language spoken in home (% English)
- Mother's education (% < HS graduate)
- Tobacco makes you cool (% definitely no)
- Wear tobacco logos (% definitely no)
- Perceived availability of drugs (% at risk)
- Community opportunities for prosocial involvement (% protected)
- School rewards for prosocial involvement (% protected)

The results indicated that the under-representation of alternative schools and student non-response bias did not affect the results of the survey. None of the weighted and unweighted estimates differed by more than 0.9 percent, except one 10th grade measure (community opportunities for prosocial involvement), which differed by 1.5 percent.

Conclusion

The analyses of school characteristics indicated that participating and non-participating schools differed in school size and the percentage of drop out/status unknown students. These differences appeared to be due to the fact that alternative schools were less likely to participate in the Healthy Youth Survey, compared with non-alternative schools. After controlling for alternative school status, these differences in school size and the percentage of

drop out/status unknown students were not significant. Based on these results, it is likely that school differences were due to the low alternative school participation.

In addition to lower participation rates of alternative schools, students in participating alternative schools had lower response rates than students in regular schools. However, analyses examining potential effects of these patterns found that under-representation of alternative schools and low student response rate in alternative schools did not appear to affect the results of the survey.

Optional "tear-off" questions

The Healthy Youth Survey contains a large number of questions, and it is well-understood that students with more difficulty reading will have trouble completing the questionnaire. Thus, data collected from questions that occur progressively later in the survey must be treated with increasing caution.

The specific purpose of this analysis was to find whether any bias was present for students who dropped out of the survey at the point of the tear-off (who attend schools that did not administer the tear-off form) in excess of bias that already exists due to non-completion as a result of reading difficulty.

Data from Healthy Youth Survey forms B and C (6th, 8th, 10th, and 12th grade) were used for this analysis. Students who provided a valid answer to the first question of the tear-off form were compared to those that did not complete the first question of the tear-off form. In order to examine bias due to the tear-off, and not bias due to non-completion, only those students who provided a valid answer to the last question on the main survey were included in the analysis.

Groups were compared by demographic characteristics that might be associated with different types of schools. These demographic characteristics include race/ethnicity, mother's education (8th, 10th, and 12th grade students only), average grades student gets in school, and smoking one or more cigarettes within the past 30 days. Comparisons were stratified by grade.

Race/ethnicity (Asian, American Indian/Alaska Native, Black, Hispanic, Pacific Islander or Hawaiian, White, other, more than one): Within 8th grade, American Indian/Alaska Native students were more likely to have received the tear-off form, compared with White students. Among 10th and 12th grade students, Asian students were more likely to receive the tear-off compared with White students. There were no differences among 6th grade students.

Mother's education ("none", "high school", "any college"). No differences among 8th, 10th and 12th grade students.

Average grades students get in school (As, Bs, Cs, and Ds/Fs): 8th grade students who reported lower grades were more likely to receive the tear-off form. There were no differences between receiving the tear-off and academic achievement for 6th, 10th, and 12th grade students.

Cigarette smoking (smoked within past 30 days, did not smoke within past 30 days): Among 8th graders, those who report cigarette smoking were more likely to receive the tear-off form. There were no associations between smoking and the tear-off form for 6th, 10th, and 12th grade students.

We did not identify consistent differences between students who completed the tear-off compared with those who did not complete the tear off, among students who completed the main body of the survey. Based on these results, there does not appear to be bias inherent in findings from the tear-off form questions, other than bias that already exists as a result of non-completion due to reading difficulty. Thus, we recommend that data collected from questions included in the tear-off should be considered of equal importance as other questions, within the limitations of the smaller numbers of participants completing the items and their later placement in the survey.

Failure to Complete Survey

Towards the end of the survey the percentage of survey questions with missing answers increased. These items, some of which were completed by fewer than 85 percent of the participants, might be subject to bias due to differences between students who were able to complete the survey in the time allotted and those who were not.

In order to assess this potential bias, students missing any of the last 30 items in Form C or before the tear-off sheet on Form B were compared to students who answered all of these questions.

Analysis was completed for grades 6, 8, 10, and 12. The following items were used for comparison:

- Language spoken in home (percentage English)
- Grade average lower than B
- Smoked cigarettes in the past 30 days
- Drank alcohol in the past 30 days
- Feel safe at school some of the time or never

[Back to Top](#)

Results

Grade	%	
6	27.1	
8	28.5	
10	20.7	
12	14.4	
All Grades	22.0	

Younger students were more likely than older students to be missing responses for any of the last 30 questions:

Fourteen out of the 20 comparisons achieved statistical significance, which is more than would be expected by chance. Significant differences included:

- 6th graders who completed Form C were more likely to have grades of B or higher, be non-smokers, and feel safe in school compared with non-completers.
- 8th graders completing Form B were more likely to speak English in the home, have grade averages of B or higher, and feel safe in school compared with non-completers.

- 10th graders completing Form B were more likely to speak English in the home, to have grade averages of B or higher, be non-smokers, and feel safe in school compared with non-completers.
- 12th graders completing Form B were more likely to speak English in the home, to have grade averages of B or higher, be non-smokers, and feel safe in school compared with non-completers.

We advise caution in interpreting the following items, administered at the end of Form B, for which there were greater than 15 percent non-completion rates:

Grade 8: D42, D44-49, P34, D13, D51, D53, D54, D16, D56, G13

Grades 6, 10 and 12 all had response rates greater than 85 percent for the last 30 questions of the survey (Table 2 and 3). Even though those who did not complete all of the last thirty questions of the survey were found to be different than those who did completed the last 30 questions (see previous analyses), the proportion of those responding (> 85 percent) should be enough to assure generalizability.

Table 2. Non- Response Rates for Last 15 Questions Grades 8, 10 and 12, HYS Form B

Question	8th	10th	12th
D42	15.2%	11.2%	6.8%
D44	16.0%	11.6%	7.0%
D45	17.0%	12.0%	7.3%
D46	17.4%	12.4%	7.3%
D47	17.8%	12.4%	7.3%
D48	18.5%	12.8%	7.7%
D49	18.3%	12.9%	7.8%
P34	18.6%	13.0%	8.0%
D13	19.6%	13.4%	8.2%
D51	19.7%	13.6%	8.2%
D53	20.4%	13.9%	8.5%
D54	20.8%	14.0%	8.6%
D16	20.9%	14.1%	9.0%
D56	21.3%	14.4%	8.8%
G13	20.9%	14.6%	9.0%

Table 3. Non-Response Rate for Last 15 Questions, Grade 6, HYS Form C

Question	6th
P04	10.0%
P33	9.4%
P34	9.6%
P35	10.5%
P36	10.5%
D34	11.1%
H33	11.3%
D54	12.3%
D45	12.5%
D46	12.8%
D42	13.4%
H03	13.1%
H06	13.9%
H09	13.0%
G13	13.5%

¹ Three separate samples were drawn for the survey: grades 6, 8 and 10/12. Grades 10 and 12 were combined for sampling because most high schools have both. However, because two schools had 10th grade only and four had 12th grade only, the participation rates for schools containing grades 10 and 12 differed slightly. Participation by grade also occasionally varied within the same school.

² Total response rates were calculated by dividing the number of completed surveys and the number of valid surveys by the number of students in all schools recruited for the state sample, based on 2004 enrollment figures.

³ All data, except for ITBS scores, were obtained from OSPI web pages:
<http://www.k12.wa.us/dataadmin/> and <http://reportcard.ospi.k12.wa.us/DataDownload.aspx>.

⁴ It is unclear why the total participate rate was higher for 6th grade alternative school compared with 6th grade regular schools. One reason may be that 6th grade alternative schools vary in their philosophy from regular elementary schools, whereas alternative high schools may work with more high-risk youth.

⁵ Total response rates were calculated by dividing the number of valid surveys by the number of students in alternative and regular schools recruited for the state sample, based on 2004 enrollment figures.

⁶ Median enrollment, as opposed to mean enrollment, was used to examine differences due to a number of outliers in the data set.