DOCUMENT RESUME

ED 234 298 CG 016 892

AUTHOR	Mauss, Armand L.; And Others
TITLE	A Longitudinal Evaluation of the "Here's Looking at You" Alcohol Education Program 1978-1981. Final Report.
INSTITUTION	Washington State Univ., Pullman. Social Research Center.
SPONS AGENCY	National Inst. on Alcohol Abuse and Alcoholism (DHHS), Rockville, Md.
PUB DATE	Nov 81
NOTE	46p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Age Differences; *Alcohol Education; Attitude Change; Decision Making Skills; Drinking; Elementary School Students; Elementary Secondary Education; High School Students; Junior High School Students; Longitudinal Studies; Models; *Outcomes of Education; *Prevention; *Program Effectiveness; Program Evaluation; *Student Attitudes; Student Behavior
IDENTIFIERS	*Heres Looking at You

ABSTRACT

"Here's Looking at You" (HLAY) is a creative prevention curriculum designed to cultivate a commitment among young people to deal in responsible ways with alcohol in their environment. The model, developed as a demonstration project for the National Institute on Alcohol Abuse and Alcoholism, consists of self-contained teaching units for grades K-12. This non-technical report, directed to non-research professionals, describes an extensive 3-year evaluation conducted to measure the effectiveness of the program in grades 4-12. Three Seattle school districts and two Portland school districts participated. Some schools were assigned to the experimental condition, and comparable schools were assigned to a control condition. The test instruments were administered to some students both before and after exposure to the HLAY curriculum; many students were followed for 3 years. Immediate results showed a definite curriculum impact for all students in the area of knowledge about alcohol and alcoholism. In addition, self-esteem was favorably affected for those in grades 5, 6, and 7. Decision making skills were improved for students in grades 6-12. Attitudes were least affected, although there was evidence that students in grade 8 and younger students changed somewhat toward favoring moderate drinking as opposed to excessive drinking. The program appeared most effective in intermediate grades, arguing for early intervention. Other influences, particularly parents, peers, and religion appeared to have a greater impact. The report contains a number of graphs illustrating the evaluation process and the results of the study. (More complex data is contained in the Scientific Appendix and the Manual of Evaluation Guidelines which are available from the Social Research Center, Washington State University.) (JAC)

A Longitudinal Evaluation of the "HERE'S LOOKING AT YOU" Alcohol Education Program 1978 - 1981

A FINAL REPORT

November, 1981

- by Armand L. Mauss, Ph.D.
Ronald H. Hopkins, Ph.D.
Ralph A. Weisheit, Ph.D.
Kathleen A. Kearney, Ph.D.

Social Research Center
Washington State University
Pullman, Washington 99164

PREFACE

It is difficult to prepare a project report that will meet the needs of the varieties of potential readers. Program and curriculum planners, teachers, and others contemplating the adoption of this kind of prevention program for their young people need to know, as simply and quickly as possible, "the bottom line" of what the evaluators have learned: Does the program in question work? To what extent, and in what ways, has the program proved capable of achieving its intended objectives? The lengthy, complex, and technical process by which the answers to such questions were obtained by the researchers is understandably a matter of secondary interest to non-researchers. The present report is prepared with this realization.

Those among our readers who want to know, in a period of ten minutes or less, our general findings about the impact of the "Here's Looking at You" Alcohol Education Program should skip immediately to the CONCLUSION (pp. 36 - 39). Those who want more information about the basis for our conclusions, still in fairly simple form, however, will wish to read some or all of the rest of the report. It is intended for non-research professionals in school systems, in community prevention projects, or in other organizations and agencies charged with finding modes of intervention to prevent or reduce the incidence and severity of alcohol-abuse among school-age children. Since the report deals only superficially with the research methodology employed, or the statistical analyses lying behind our findings, the reader is, in effect, being asked to assume that the evaluators did their work competently and objectively. A relatively short and simple report like this one, however, may be somewhat deceptive, for it cannot possibly convey the complexities, the time, or the anguish of the three years' labor from which it derives.

Readers who would like to know more about such matters are invited to send for certain other documents from our project files. First of all, there is the Scientific Appendix to this report, which may be obtained by sending a written request and twenty dollars ($20.00) to the Social Research Center, Washington State University, Pullman, WA. 99164. This Appendix covers the questionnaire items comprising our instruments, and the frequency distributions for each item; how we constructed our scales and other variables; descriptive statistics (means, standard deviations, etc.) for our major variables, both in our cross-sectional and in our longitudinal samples; the basic comparisons lying behind our assertions about curriculum impact; the theoretical model and analytic strategies guiding our ongoing analyses; and certain other matters of scientific interest. Another document of potential interest to readers of this brief report is our Manual of Evaluation Guidelines for the "Here's Looking at You" Alcohol Education Program, nicknamed our "cookbook." This document can also be obtained from the same address for another $20.00. It is intended for on-site evaluators with at least some experience, but not necessarily with any statistical sophistication. The "cookbook" is a step-by-step guide covering all the evaluation procedures for the program on a 1-year (or year-to-year) basis, without any longitudinal component. It begins by discussing arrangements with schools and parents for implementing and evaluating the program, and then goes into data collection planning, procedures, and instruments, the coding of the instruments, and the basic procedures and formats for analyzing and reporting the data. Copies of the test instruments for each grade level are included, along with other samples

or facsimilies of evaluation materials. Still other documents from our files that may be of interest to our readers are listed in the Bibliography at the end of this report.

Those who wish to know more about the "Here's Looking at You" program itself, its philosophy, materials, teaching methods, or teacher training, should not write to us but instead to Roberts and Associates, 9131 California Avenue, SW, Seattle, WA 98136. (Phone: 206-932-8409.) Our work has to do only with evaluation research, and those who want information, beyond what is contained in this report, about the methodological, technical, or statistical details of our work should write to us for the Scientific Appendix, the Manual of Evaluation Guidelines, and/or other documents listed in our Bibliography.

A. L. M.
R. H. H.
R. A. W.
K. A. K.

Pullman, Washington
November, 1981.

TABLE OF CONTENTS

INTRODUCTION

Background

Our relationship as evaluators to the "Here's Looking at You" Alcohol Education Program (hereinafter HLAY Program) goes back to the late Fall of 1975. Earlier that year, the Health Education Department of the Educational Service District No. 121, Seattle, Washington, had received a 3-year grant from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) to create and field-test a model alcohol education program as one of several NIAAA demonstration projects. The principal investigator and creator of the HLAY program at ESD-121 was H. Clay Roberts, and chief among his several collaborators on the project was Carol Mooney. They produced what is, in our opinion, a highly creative and imaginative prevention curriculum derived from the most advanced theoretical assumptions current among health educators during the 1970s. The ultimate objective of the curriculum is to cultivate a commitment among young people to responsible ways of dealing with alcohol in their environment. This includes their decisions in alcohol-related situations, such as riding with someone who has been drinking, as well as their personal drinking behavior. In this context, "responsible" decisions would include both abstinence and light-to-moderate drinking, depending upon the values of the person, but would exclude heavy drinking, drunkenness, and other choices which might lead to anti-social or self-destructive consequences.

In accordance with the terms of their NIAAA grant, Mr. Roberts and his colleagues at the ESD called for bids on an evaluation contract, and at length concluded one with us at Washington State University (WSU), after the HLAY program had been created and was ready for implementation. That contract and its renewals lasted until the Spring of 1978. It required us first to evaluate the teacher-training component of the HLAY program, in two phases, between which certain modifications and "fine-tuning" were carried out by the ESD educators to enhance the effectiveness and the efficiency of the teacher training. Also evaluated was the usefulness of the "multiplier effect" in teacher-training--that is, the reliance on ESD-trained teachers to train fellow teachers back in their respective school buildings. Finally, under our 1975-1978 contract with ESD-121, we created and pilot-tested a research design, and some research instruments, for evaluating the HLAY program in the classroom for grades 4 through 12. This was a minimal research effort intended to get preliminary information about actual teacher performance, as well as about the impact of the new program on the students themselves. Additional information about these early evaluation efforts and outcomes may be obtained from the publications and project reports listed here in our bibliography under Rankin, et al., 1978; Tarnai, et al., 1978; Mauss, et al., 1980; and Tarnai, et al., 1981.

The next phase of our work, from 1978-1981, is what is covered by the present report. In mid-1978, we at WSU received a grant from NIAAA to conduct an extensive and thorough 3-year evaluation of the HLAY alcohol education program. This new grant called for a kind of reversal in our relationship with the program developers at the ESD-121: We, and not they were to carry the initiative in the decisions about the dissemination and evaluation of the program. These decisions would be based upon our quasi-experimental research design, which designated those schools and classrooms at the various grade-levels that would serve as experimental or as control groups. Our ESD-121 partners would train the teachers and disseminate the teaching materials in the schools that we

targeted. Their partnership and consultation in the research endeavor were obviously essential, so they too were given a new NIAAA grant in mid-1978 to work with us. The entire enterprise, however, was now defined as primarily a research project, rather than as a program demonstration project.

The Nature of the Product Being Evaluated

As we observed above, the Here's Looking at You" Alcohol Education Program is aimed at cultivating "responsible" ways of dealing with alcohol on the parts of young people. As far as personal behavior is concerned, for some of the youth, "responsible" in this context may mean total abstinence, depending upon religious or other personal values. For those who choose to drink alcoholic beverages, however, the operational meaning of "responsible" will be a mode of decision-making and of drinking behavior that is free of problematic consequences or other indications of abuse. Responsible decisions and demeanor will also express themselves in appropriate efforts to discourage abuse among friends and others. Implicit in the theoretical basis of the HLAY curriculum is the assumption that abusive or problematic drinking behavior is attributable to many factors, including alcoholism; ignorance about the drug alcohol and its physiological or psychological functions; low self-esteem; difficulties in using decision-making skills; and difficulties in identifying and applying appropriate coping strategies. Accordingly, the curriculum contains components aimed at enhancing knowledge, self-esteem, and skills in coping and in decision-making. The ultimate goal of the entire alcohol-education package, of course, is to impact actual behavior where alcohol is concerned.

To meet its cognitive, affective, and behavioral objectives, the HLAY program has been packaged in self-contained teaching units for each grade level, kindergarten through high school. These units can either be taught on successive days, which will take about three school weeks, or they can be spread out and integrated among the existing curricula in health education, in biological science, in social studies, in home economics, or in nearly any other traditional subject-matter area. All of the necessary teaching materials for the curriculum have been prepared and boxed up in specific grade-level kits, which include games, films, and visual aids of all kinds, as well as the substantive teaching guides and textual materials. Three full days of teacher training are normally required for adequate preparation in the classroom use of these materials. Readers who are interested in knowing more about the "Here's Looking at You" alcohol education program, its kits and materials, the teacher training, and so on, should make direct contact with Roberts and Associates (see Preface).

THE EVALUATION DESIGN AND DATA

The research design for our evaluation, and the nature of the data we collected, if treated thoroughly here, would require a lengthy and complex discussion that we assume would hold little interest for most of our readers. Those scientific colleagues in the research community who would like the benefit of such detailed information are invited to send for the Scientific Appendix to this report, which can be purchased separately, as we have explained in the Preface.

The Basic Design

For the benefit of the general professional reader, we will explain briefly that we used what program evaluators usually call a "quasi-experimental" design. We assigned some schools to an "experimental" condition and some comparable schools to a "control" condition; or, as an occasional variation, some classrooms within a school to the experimental condition, and other classrooms within the same school to a control condition. Our tests or questionnaires, designed to measure the effectiveness of the HLAY program, were administered to the "experimental" students after they had been exposed to the curriculum for a given school year, while the same instruments were administered to the "control" students without their having been exposed, or before they were exposed. Some of our experimental students were deliberately administered the test instruments both before and after exposure to the HLAY curriculum (pre-test and post-test) to allow us to assess the actual a priori comparability of our experimental and control groups.

Altogether, five different school districts participated in our evaluation effort: three in the Seattle area and two in the Portland area. Some of these districts were large and varied enough to provide us with both experimental and control students; others provided experimental students only, or else experimental students at some grade levels and control students at others. We deliberately chose school districts that would provide us with as much variety as possible across the entire project, including small and large districts, rural and suburban districts, urban schools with and without much ethnic variety, and so on. Naturally, we had to take the students, classrooms, and schools as we found them already constituted by the district administrations, so we were not able to assign the students randomly to the experimental or the control conditions. Nevertheless, we did succeed in negotiating with each school district as much leeway as feasible in making random assignments of schools and/or classrooms, and we sought for as much similarity or comparability as possible between our control and experimental schools and rooms.

A particularly valuable feature of our evaluation was its longitudinal component, in which we followed many of our control and experimental students across all three years. Depending upon when certain students entered or left our project, and upon curriculum variations at certain grade-levels or schools, some of our experimental students were exposed to the HLAY program only one of the three years, some were exposed two years, and some all three years. A one-time exposure could have occurred, of course, in any of the three years. A two-time exposure could have occurred either in years one and two, in years two and three, or in years one and three. Whether exposed or not, students whom we tested in any given year were followed up whenever possible and tested during the subsequent years of the project as well. Our research design, interacting as it did with the many natural contingencies and changes in the school populations across time, was enormously complex, a situation compounded by the use of a system of self-generated private identification codes, designed to protect the anonymity and confidentiality of student responses on our test instruments (Kearney, 1982). Nevertheless, all of these complexities had to be accepted, for they enabled us to assess the impact of the HLAY curriculum not only in the immediate sense, after only one year's exposure, but also in the longer run and after varying degrees and modes of cumulative exposure. Furthermore,

any HLAY impact on actual drinking behavior could be assessed only after some lapse of time beyond the exposure itself.

The Nature, Quality, and Quantity of the Data

The data for evaluating the HLAY program were collected in the classrooms by the teachers with tests provided by our research project. One period, usually 40 or 45 minutes, was required for administering a test in each classroom. Most test items were of the multi-choice type, though a few required students to fill in blank spaces. All teachers were first trained in the proper procedures we had devised for administering the tests, and shown how to create the testing conditions that would inspire the trust of the students in our guarantees of anonymity. Teachers of experimental students received this test training during the regular 3-day training workshop called for in the HLAY program. Teachers of control students received their test training in special sessions before school on or near the day of testing and were in addition paid modest fees for doing our data collection, if they and their students were not otherwise eventually to benefit by the HLAY program. Even after the training, a schedule of techniques for follow-up and supervision was followed, to insure as much compliance by the teachers as possible with our standardized testing procedures.

The same tests (questionnaires) were used as the instruments for gathering our data throughout the entire project, regardless of district, school, or assigned condition (control, experimental, pre- or post-test). The tests did differ, however, by grade-level: the tests for grade 4, grade 5, grade 6, junior high and senior high classes were constructed and field tested (pilot-tested) separately. (No testing was attempted below the 4th grade because of the expense and uncertainty involved in trying to create the non-pencil/paper tests required below that reading level.) The test-retest reliability coefficients calculated after pilot-testing ranged mostly from .50 to .90 for the various items in our instruments. Reliability coefficients falling below that range tended to occur with the younger children and with attitude measures, which seemed the least stable. Among the most stable (reliable) were the self-reports of drinking and drug-using behavior. Where we constructed composite measures based upon more than one test item, the alpha reliability coefficients were the larger, of course, the more the test items that were included. The construct validity of the test items relating to the HLAY curriculum was usually established through verification by the health educators at ESD-121. The validity of the test items not related to the curriculum was usually established either internally or by reference to their successful uses elsewhere in the professional literature. (More on validity and reliability can be found in either the Manual or the Appendix mentioned in the Preface to this report.)

The contents of the tests at the various grade levels depended in part upon the contents of the curriculum to which they respectively pertained at each grade level. Otherwise, they tended to have a cumulative quality from one grade-level to the next: that is, the test at any given grade level tended to include or repeat the measures that had appeared at the earlier grade levels. In the junior and senior high grades, the HLAY curriculum and accompanying tests were given at whichever specific grade-level the school administration selected: that is, either grade 7 or 8 or 9 for junior high, and either grade 10 or 11 or 12 for senior high (though sometimes grade 9 was included at the senior high level).

The factors measured by our instruments at the various grade levels, including both curriculum-related and other kinds of measures, are indicated by the summary in Table 1. Single copies of our test instruments at each grade level may be obtained free of charge by writing to the Alcohol Education Research Project, Social Research Center, Washington State University, Pullman, WA 99164.

Table 1

Variables Measured by Student Instruments*

Variable	Grade Level				
	4	5	6	JH	SH
Personal Code	x	x	x	x	x
Demographic Variables					
Gender	x	x	x	x	x
Family size	x	x	x	x	x
Birth order	x	x	x	x	x
Age	x	x	x	x	x
Ethnic group	x	x	x	x	x
Knowledge*	x	x	x	x	x
Self-Esteem*	x	x	x	x	x
Decision-Making*					
Assigning problem responsibility				x	x
Generating alternatives	x	x	x	x	x
Selecting responsible alternative	x	x	x	x	x
Assessing advantages and disadvantages	x	x	x	x	x
Attitudes about Alcohol*					
Total attitude score	x	x	x	x	x
Alcohol for mood enhancement			x	x	x
Tolerance of abstinence	x	x	x	x	x
Intolerance of abuse	x	x	x	x	x
Tolerance of moderate use	x	x	x	x	x
Influence of others			x	x	x
Alcoholism as disease, not character defect			x	x	x
Treatment of alcoholics				x	x
Childhood Drinking Behavior	x	x	x		
Adolescent Drinking Behavior					
Irresponsible uses				x	x
Problem drinking				x	x
All drugs				x	x
All drugs but alcohol				x	x
Current drinking situation				x	x
Expectations about drinking				x	x
Monthly frequency				x	x
Yearly frequency				x	x
Quantity				x	x
$QF_{monthly}$				x	x
QF_{yearly}				x	x
BAC level				x	x

Table 1 (continued)

Variables Measured by Student Instruments*

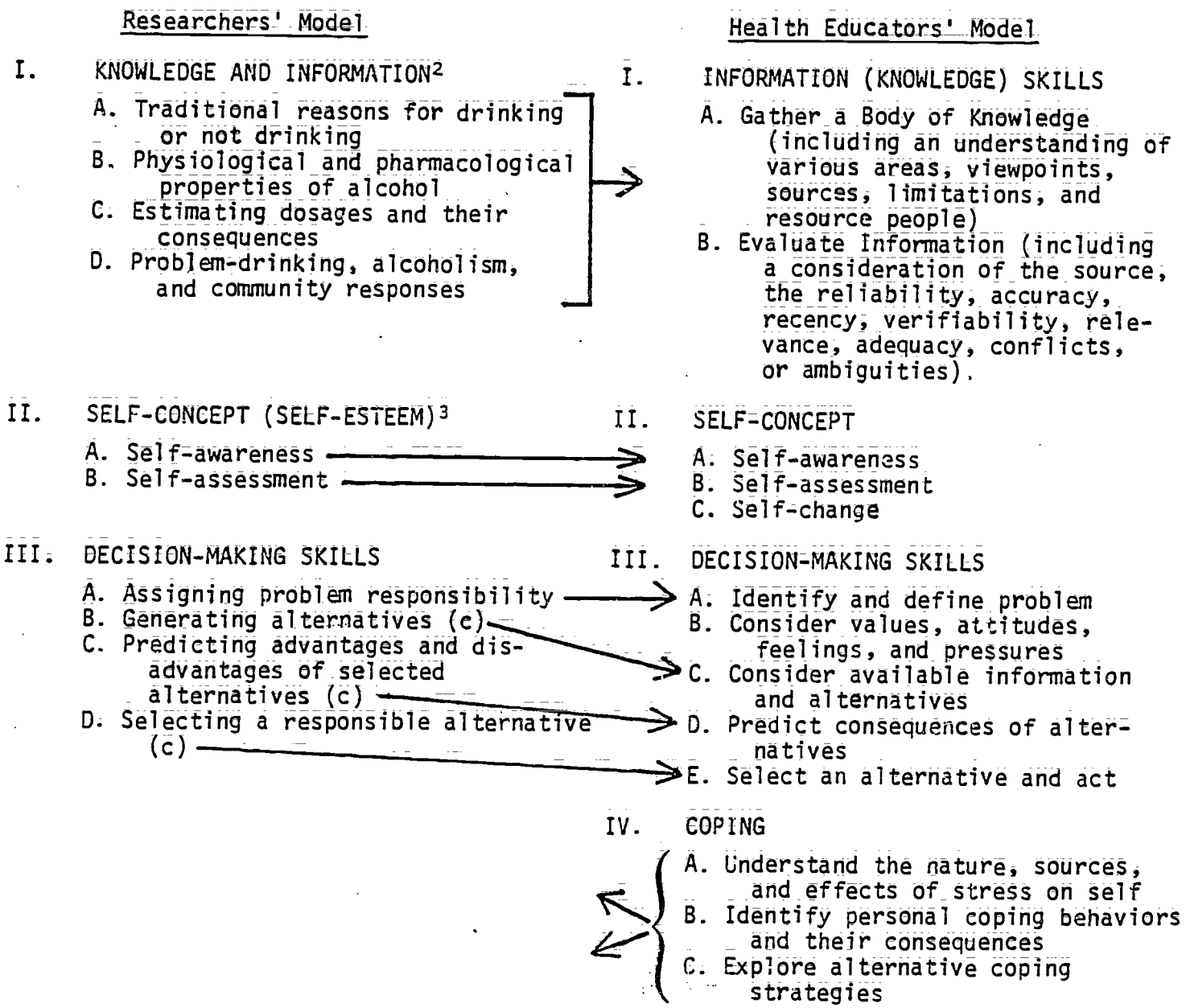
Variable	Grade Level 4	5	6	JH	SH
Attitude Toward Alcohol Education				x	x
Peer/Parental Influence					
Peer support	x	x	x	x	x
Peer control	x	x	x	x	x
Peer drinking influence				x	x
Parental support	x	x	x	x	x
Parental control	x	x	x	x	x
Peer loyalty				x	x
Parent versus peer				x	x
School Variables					
Extracurricular activities				x	x
Level of performance	x	x	x	x	x
Satisfaction	x	x	x	x	x
Other Non-Curriculum Variables					
Religious control/constraint				x	x
Conventional commitment				x	x
Smoking behavior				x	x
General health				x	x
General happiness	x	x	x	x	x

*NOTE

It is important to emphasize that this table organizes the curriculum variables according to the conceptualizations of the researchers, not necessarily of the curriculum developers themselves. The differences between the two can be attributed mainly to some evolution in the thinking of both parties after their original agreement upon the construction of the test instruments, and also to different conventions in the professional literatures, respectively, of social science research and health education. In any case, the differences between the two conceptualizations are not great: Both identify knowledge or information about alcohol and alcohol abuse, self-concept, and decision-making skills as three of the major curriculum components. The main difference lies in the fourth component, which the curriculum people identify as coping (defined as "using what one knows to feel better"). The researchers have instead focused on attitudes as the fourth major component (usually meaning dispositions toward "responsible" as opposed to "irresponsible" ways of dealing with alcohol in one's environment). The curriculum people, on the other hand, see attitudes not as a separate component but rather as permeating the entire processes of coping and of decision-making. Thus, some of the test items that might be regarded as indicators of "coping" or of "decision-making" by the health educators are classified in this table under attitudes, and others perhaps elsewhere. It should be noted also that our self-concept variable is made up essentially of indicators of self-esteem (following Coopersmith). This overlaps with, but is not identical to, the three dimensions of self-concept now identified by the curriculum developers as "self-awareness," "self-assessment," and "self-change." (See Figure 1)

Figure 1

COMPARISON AND RECONCILIATION OF THE RESEARCHERS' AND THE HEALTH EDUCATORS' CONCEPTUALIZATIONS OF THE "HERE'S LOOKING AT YOU" CURRICULUM[1]



IV. ATTITUDES

A. Total (composite) attitude score
B. Attitudes toward alcohol for mood enhancement (c)
C. Tolerance toward abstinence (d)
D. Tolerance toward moderate alcohol use (d)
E. Intolerance toward alcohol abuse (d)
F. Attitude toward influence of others (c, d)
G. Recognizing alcohol as disease, not character defect
H. Appreciation of need for treatment of alcoholics

(See next page for notes to this Figure)

NOTES to Figure 1:

[1]Under the "Researchers' Model" below is a classification of what the researchers actually attempted to measure, in accordance with the understanding they took from consultations with the health educators at ESD No. 121 at the beginning of the project. Under "Health Educators' Model" is a classification based on a more recent document provided by the ESD No. 121 people. It is important to keep in mind that this model represents how they conceive the curriculum, rather than what the researchers tried to measure. The purposes of the two models are thus somewhat different, but the differences between them are not great.

[2]Not all four of these aspects of knowledge were measured at all grade levels, and they were not kept separate at all for purposes of measurement. Rather, a composite knowledge score was computed.

[3]The researchers always conceived of self-concept primarily as self-esteem, for measurement purposes, and used a modified version of a standard self-esteem scale developed by Coopersmith. As a post-hoc conceptual effort, it is feasible to see some of the self-concept items in this scale as dealing with self-awareness, and others as dealing with self-assessment. However, in actual measurement, no such distinctions were made. A composite score was derived for the entire 10-item scale.

(c) = indicates that this measure developed by the researchers relates to some of what the health educators consider "coping."

(d) = indicates that this measure developed by the researchers relates to some of what the health educators consider "decision-making."

It will be apparent that in the health educators' model there is not a total and discrete separation between coping and decision-making, especially with regard to generating, selecting, and predicting the consequences of alternative choices. This overlap accounts in part for the reconceptualization made by the researchers for measurement purposes.

The data generated by the number and variety of testing situations explained above amounted, of course, to thousands of cases altogether. Table 2 shows the total numbers of questionnaires received (pretest, posttest, and control) across the grade levels for the three years of the project.

Table 2

Total Numbers of Questionnaires Received (Pretest, Posttest, and Control)

Proj. Year	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	Gr 9	Gr 10	Gr 11	Gr 12	Total
1978-79	628	626	696	215	660	230	72	46	120	3,293
1979-80	905	914	1037	891	1329	712	153	196	425	6,573
1980-81	800	840	901	1319	987	1042	755	380	521	7,545
TOTALS	2333	2380	2634	2425	2976	1984	980	622	1077	17,411

One useful subset of the data for descriptive purposes is that based on all "naive" (unexposed) students from project years 01 and 02. These are the questionnaires (pretests and control) from students who, before completing the questionnaire, had never been exposed to the model alcohol education curriculum or tested with our instruments. Thus, the data from these naive students permit description of the population of students as we found them, providing normative information about the distributions of the scores on the variables we measure, in the absence of any experience with the curriculum or test instruments. The numbers of cases of these "naive" students are presented in Table 3 (grades 10 and 11 have been combined to provide a respectable number of cases at that level).

Table 3

Numbers of "Naive" Pretest and Control Students from Years 01 and 02

Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	Gr 9	Gr 10-11	Gr 12	Total
1109	404	476	550	1125	493	262	430	4,849

For the assessment of immediate, short-term impact of the curriculum, another data subset was constructed consisting of the questionnaire responses from experimental and control students who were new to the project in Years 01 or 02. That is, the control data are from control students the first time they were tested, and the experimental data are from students who were exposed to the curriculum for the first time and then posttested. The resulting numbers of these students are presented in Table 4.

Table 4

Control and Posttest-Experimental Data for Assessment of Immediate, Short-Term Curriculum Impact (Years 01 and 02 Only)

	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	Gr 9	Gr 10-11	Gr 12	Total
Control	699	272	331	426	765	114	116	265	2,988
Posttest	834	366	365	215	1047	493	220	280	3,820

Of course, longitudinal samples are required for evaluating the cumulative or longer-term impact of the curriculum. The numbers of students whose questionnaires could be linked across all three years of the project (both experimental and control students) are presented in Table 5. These cases constitute the data subset for which cumulative impact has thus far been examined.

Table 5

Total Numbers of Cases of Three-Year Linked Longitudinal Data

Grade Cohorts:	4-5-6	5-6-7	6-7-8	7-8-9	8-9-10	9-10-11	10-11-12	Total
	300	266	226	13	62	30	0	897

The numbers of three-year linked cases are relatively small, of course. However, we also have longitudinal data for many more cases where the linkage was possible for only two years. The total numbers of such two-year linkages (both experimental and control) is given in Table 6, which includes the year 01-to-02 linkages plus (for those districts joining our project in year 02) the year 02-to-03 linkages. This data set will also eventually be used to examine the cumulative and longer-term impact of the curriculum. It should be noted that the cases represented in the three-year longitudinal data set (Table 5) are a subset of the data linkable across two years (Table 6).

Table 6

Total Numbers of Cases of Two-Year Linked Longitudinal Data

Grade Cohorts:	4-5	5-6	6-7	7-8	8-9	9-10	10-11	11-12	Total
	457	426	323	416	813	375	11	116	2,937

Analyses done so far on comparisons between the total data set and each subset (whether longitudinal or cross-sectional) have uncovered no significant systematic biases in the subsamples. We are therefore quite confident that our subsamples are fairly representative of the entire project. Nor, wherever the data permitted us to compare the five participating school districts with each other, did we find any consistent biases or differences by district, except for understandable differences in certain demographic traits (like race/ethnicity). The ethnic distribution across the samples, incidentally, for all districts combined, was approximately 68% Caucasian, 12% Black, 3% Indian/Native American, 2% Hispanic, 7% Asian, and 8% other or ambiguous. The samples were almost evenly divided between males and females.

THE SHORT-TERM IMPACT OF THE ALCOHOL EDUCATION CURRICULUM

Now that the reader has, we hope, a fairly clear idea of the design and the data-base for our evaluation, we are ready to review some of our results. We begin with the immediate or short-term impact of the HLAY curriculum, that impact which we could measure after only one exposure of the students to the alcohol education unit. We remind the reader that no immediate curriculum impact on actual drinking behavior could feasibly be measured, given that the test instruments were normally administered at the completion of the unit, and that the measures dealt with drinking behavior in the recent past, generally before exposure to the HLAY curriculum. Measures of immediate impact, then, are limited to the various cognitive and affective components of the curriculum: knowledge and information about alcohol, etc.; self-esteem; coping and decision-making; and attitudes.

Table 7 summarizes the results of literally months of careful comparisons and analyses. The various components and sub-components of the HLAY curriculum at the respective grade levels are listed down the left side of the table, and plus signs (+) indicate at which grade levels a curriculum impact was established unambiguously and with statistical significance ($p = .05$ or less). Minus signs indicate the few places where curriculum exposure was followed by the opposite of the expected results. The large number of remaining (blank) spaces indicate where the curriculum impact was ambiguous or not statistically significant, even though many times it was in the expected direction. Standards for the determination of unambiguous impact (+) were quite rigorous, so these results are probably conservative estimates of impact. Judgments were based not only on sheer differences between experimental and control students, or on pre-test vs. post-test differences, but also on considerations of comparability between control and experimental samples, test effects (or sensitizing) on post-tests from pre-testing, and other such contingencies explained in our Scientific Appendix. The results in Table 7 cover all five of our project districts combined; as we mentioned above, our design was not equally elaborate in all districts, but we found few systematic individual district differences or idiosyncrasies in our data.

The bar graphs (Graphs 1 through 5) also provide selected illustrations of some of the comparisons that lie behind Table 7. These particular illustrations were selected here because they are among the most positive examples of curriculum impact. In general, the table and the illustrative graphs are quite encouraging in what they indicate about immediate impact. There are more unambiguous instances of impact for the various curriculum components and the various grade levels than we have ever been able to establish heretofore, either in earlier (and more tentative) analyses of these data, or in the pre-1978 period of our work (see Tarnai, et al., 1978; and Mauss, et al., 1980). We have nearly always found an immediate impact on the knowledge/information component, but the more affective components have always proved much more resistant to successful outcome and/or measurement than is the case here. The HLAY has clearly enhanced self-esteem after only one exposure at grades 5, 6, and 7, at least. Various aspects of decision-making and coping skills have also been improved, especially from grade 6 on up. Attitudes toward alcohol use have always proved the hardest to change in the short run, but at least there is some evidence here that students exposed to the HLAY program in grades 4, 6, 7, and 8 were changed toward a posture of moderate use after one curriculum exposure. Implications of all this for actual behavior toward alcohol would presumably be favorable, but we will need to turn to our longitudinal data to look at that issue.

Table 7

Summary Evidence of Unambiguous Immediate Curriculum Impact

	Grade Level							
	4	5	6	7	8	9	10-11	12
Knowledge	+	+	+	+	+	+	+	+
Self-Esteem		+	+	+				
Decision-Making								
Assigning Problem Responsibility				+	+	+	+	
[1]Generating Alternatives		-	+	+	+		+	
[1]Selecting Responsible Alternative						+	+	
[1]Advantages and Disadvantages			+		+			-
Attitudes about Alcohol								
Total Attitude Score (all dimensions)								
[1]Mood Enhancement				-	-			
[2]Tolerance of Abstinence								
[2]Intolerance of Abuse								
[2]Tolerance of Moderate Use			+	+	+			
[1]Influence of Others				-				
Alcoholism as Disease, not Character Defect								
Treatment of Alcoholics								

[1]Considered also an aspect of coping by curriculum developers. (See footnote, Table 1)

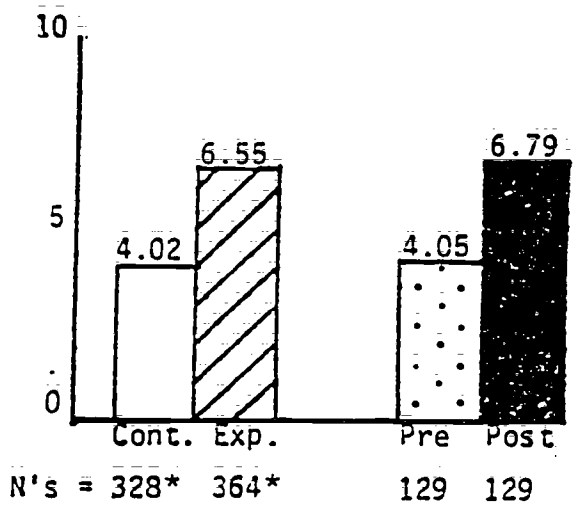
[2]Considered also an aspect of decision-making by curriculum developers.

GRAPH 1

A. Immediate Curriculum Impact on

KNOWLEDGE

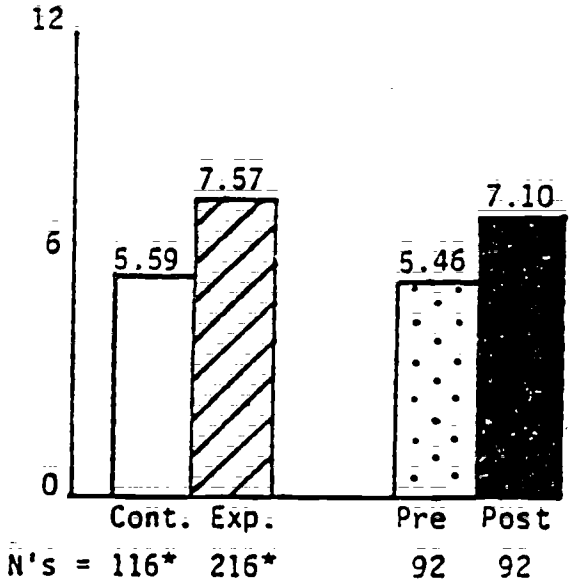
Grade 6



B. Immediate Curriculum Impact on

KNOWLEDGE

Grades 10 & 11



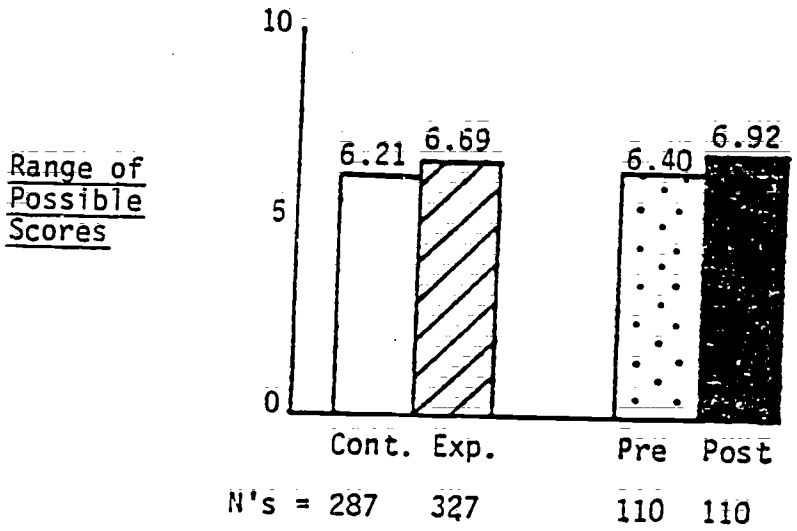
*These Ns correspond to their counterparts in Table 4, except for minor variations resulting from missing data (non-responses) on specific items

GRAPH 2

A. Immediate Curriculum Impact on

SELF-ESTEEM

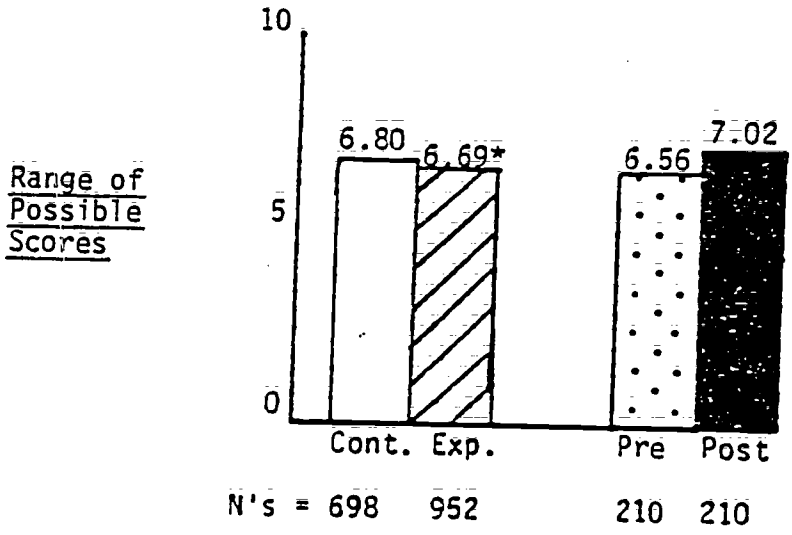
Grade 6



B. Immediate Curriculum Impact on

SELF-ESTEEM

Grade 8



*Not statistically significant

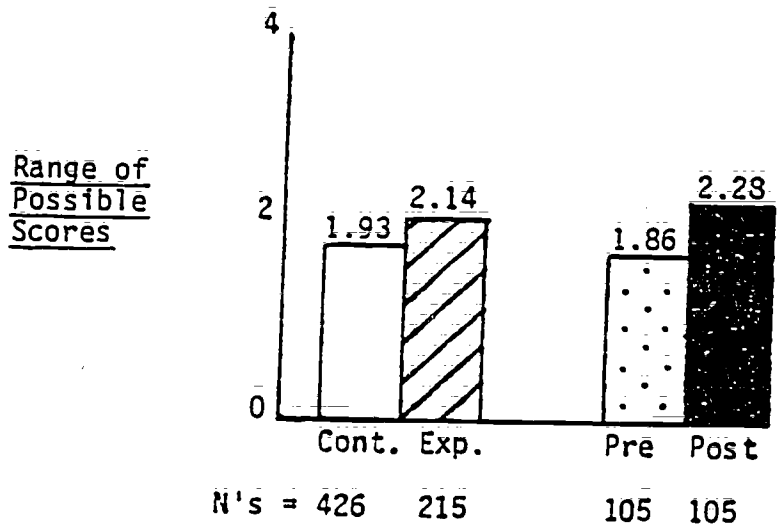
GRAPH 3

A. Immediate Curriculum Impact on Decision-Making Skills:

ASSIGNING PROBLEM RESPONSIBILITY

(Not included in curriculum before Grade 7)

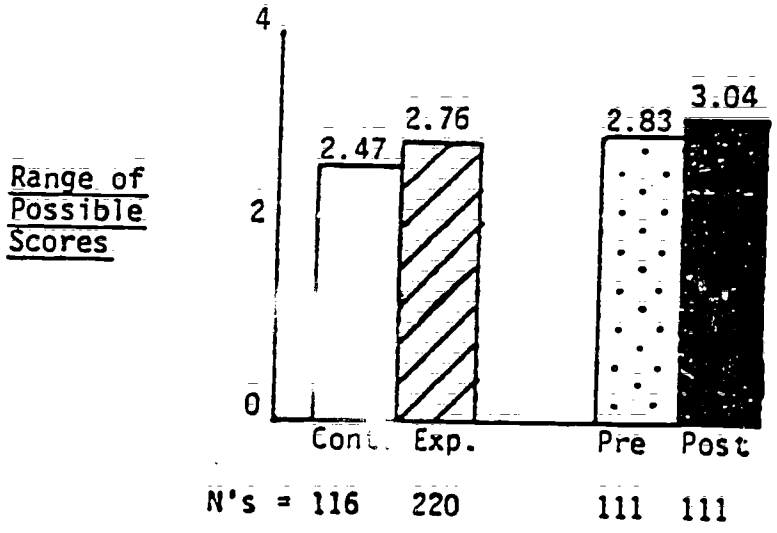
Grade 7



B. Immediate Curriculum Impact on Decision-Making Skills:

ASSIGNING PROBLEM RESPONSIBILITY

Grades 10 & 11

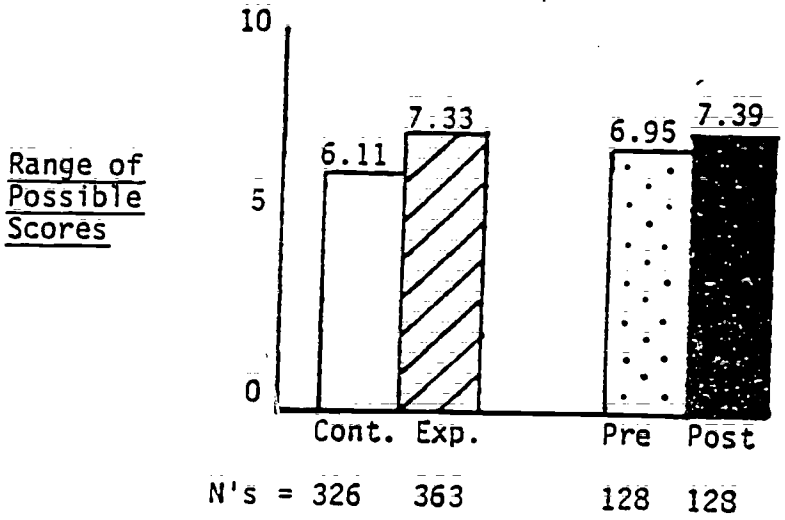


GRAPH 4

A. Immediate Curriculum Impact on Decision-Making Skills:

GENERATING ALTERNATIVES

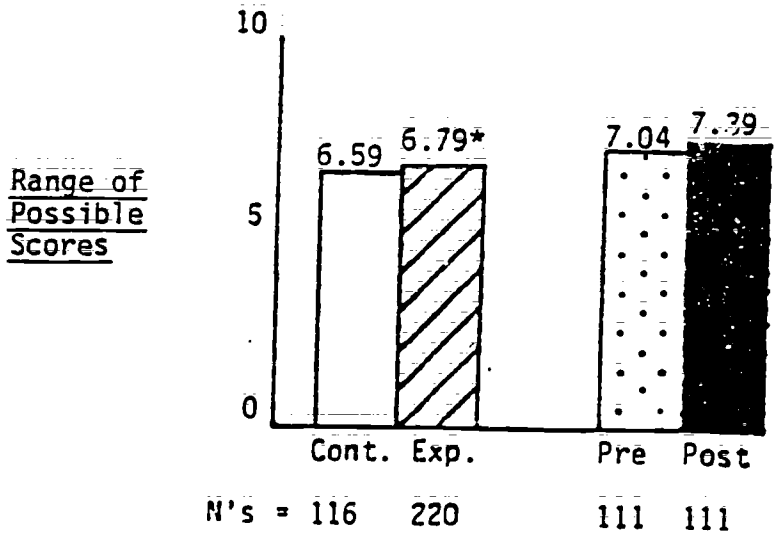
Grade 6



B. Immediate Curriculum Impact on Decision-Making Skills:

GENERATING ALTERNATIVES

Grades 10 & 11



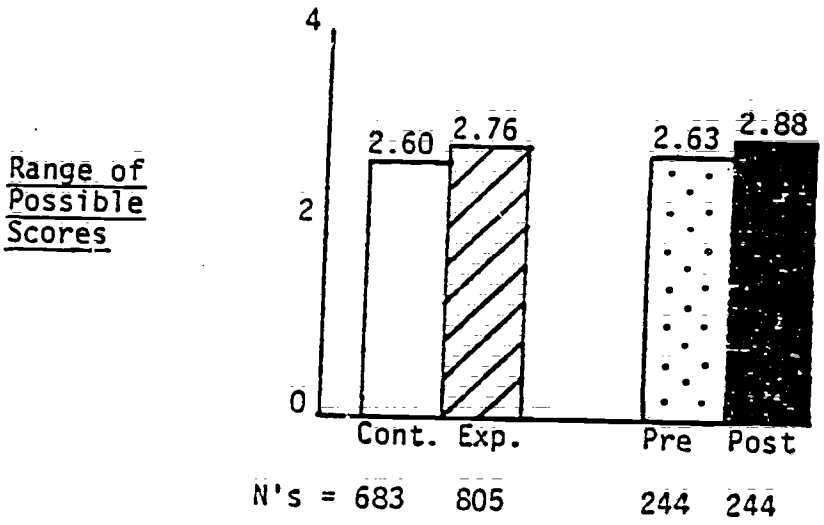
*Not statistically significant

GRAPH 5

A. Immediate Curriculum Impact on Attitudes: *

TOLERANCE FOR MODERATE ALCOHOL USE

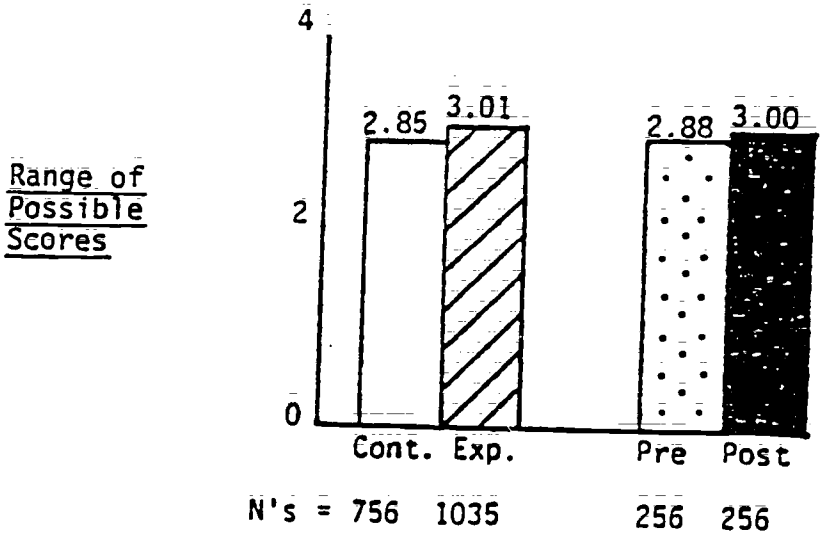
Grade 4



B. Immediate Curriculum Impact on Attitudes:*

TOLERANCE FOR MODERATE ALCOHOL USE

Grade 8



*Considered by curriculum developers as an aspect of the decision-making process

CUMULATIVE AND LONG-TERM IMPACT OF THE CURRICULUM

As we explained above, we have data sets that will allow us to assess the impact of the HLAY curriculum either over a 2-year span or over a 3-year span, or both. Our analysis so far has been limited to the 3-year data, and we are assuming for purposes of this report that cumulative curriculum impact not appearing in three years would not have appeared in two. With measurements taken after a lapse of three years, we also have logical grounds for drawing some conclusions about eventual impact on actual drinking behavior, as well as continuing impact on the other (more cognitive and affective) student characteristics addressed by the HLAY curriculum.

Table 8 provides a summary of curriculum impact sustained over three years in three age cohorts: one which started in the fourth grade in the first year of our project and ended in the sixth grade by our third year (here called the 4-5-6 cohort); one which progressed through grades 5, 6, and 7 during the life of our project; and one which moved through grades 6, 7, and 8. We were not able to gain access to enough cases of 3-year linked data for any one cohort ending above the eighth grade to make feasible the analysis of curriculum impact at those higher grades. This outcome is partly the result of the relatively rare occurrence of repeated HLAY program exposure above the eighth grade; that is, the high schools in our project generally installed the HLAY program in their curricula in <u>either</u> grade 9 <u>or</u> grade 10 <u>or</u> one of the higher grades, but not in <u>more</u> than one of the high school grades. However, since we have learned that the attitudes and behavior of young people toward alcohol are likely to be well established by the end of the junior high period, our data do cover probably the most formative crucial years.

The plus signs (+) on Table 8 indicate where the three cohorts respectively "ended up" with regard to knowledge and information about alcohol, self-esteem, decision-making skills, attitudes, and behavior. Again depending upon how many HLAY exposures the various curricula called for during the middle-school or junior high years, the experimental students in these cohorts had differential patterns of exposure: the 4-5-6 cohort was exposed three times (once in each of the three project years); the 5-6-7 cohort was exposed at least the first two years and <u>some</u> of them all three years; and the 6-7-8 cohort was exposed <u>either</u> in years one and three (an XOX pattern) <u>or</u> in years two and three (an OXX pattern). All students, whether experimental or control, in all three cohorts, were measured (that is, tested) all three years. The plus signs indicate statistically significant differences between the experimental and control students that occurred (and remained) in the cohort <u>in or by the third year</u>. On some of the curriculum dimensions, such as knowledge, there was evidence of program impact in all three years. On some of the other dimensions, it occurred during the second year and remained into the third. On still others, the impact appeared (or became statistically significant) only in the third year.

In general, Table 8 is not as encouraging as was the earlier table on immediate curriculum impact. Aside from the knowledge dimension, which almost routinely yields to instruction, the other plusses on Table 8 are few and far between. Where the 5-6-7 cohort is concerned, there are no more plusses at all, which may be partly attributable to the fact that some of the measuring devices used in the instruments (tests) are somewhat different, starting in grade 7, than they had been in the earlier grades. On the whole, though, we are inclined

TABLE 8

SUMMARY EVIDENCE OF CUMULATIVE CURRICULUM IMPACT

BASED ON THREE-YEAR LONGITUDINAL SAMPLES

	Cohort 4-5-6	Cohort 5-6-7	Cohort 6-7-8
Knowledge	+	+	+
Self-Esteem	—	—	+
Decision-Making			
Assigning Problem Responsibility			
[1]Generating Alternatives	+	—	—
[1]Selecting Responsible Alternative	—	—	—
[1]Advantages and Disadvantages	+	—	+
	—	—	+
Attitudes About Alcohol			
Total Attitude Score (all dimensions)	—	—	—
[1]Mood Enhancement	—	—	—
[2]Tolerance of Abstinence	—	—	—
[2]Intolerance of Abuse	—	—	—
[2]Tolerance of Moderate Use	—	—	—
[1]Influence of Others	—	—	—
Alcoholism as disease, not character defect	—	—	—
Treatment of Alcoholics		—	—
		—	—
Drinking Behavior			
Irresponsible Uses			
Problem Drinking		—	—
All Drugs		—	+
All Drugs but Alcohol		—	—
Current Drinking Situation		—	—
Monthly Frequency		—	—
Yearly Frequency		—	—
Quantity		—	—
QF Monthly		—	—
QF Yearly		—	—
BAC Level		—	—
		—	—

[1]Considered also an aspect of coping by curriculum developers (See footnote, Table 1).

[2]Considered also an aspect of decision-making by curriculum developers.

to think that if our instruments were sensitive enough to pick up so many more instances of curriculum impact in the short run (that is, in Table 7), then they should have picked them up in the longer run, too, if they had occurred.

In the case of the attitudes dimension of the HLAY program, no enduring curriculum impact was observable, but there were instances of short-term impact, both in these 3-year data and in the cross-sectional data in Table 7. On the decision-making dimension, long-term effects of the curriculum were clear for some of the skills in the 4-5-6 cohort and again in the 6-7-8 cohort. Perhaps most encouraging, however, was the evidence of impact on problem-drinking in the 6-7-8 cohort. This actually appeared in year two for this cohort and was sustained into year three. Since these junior high years were the ones in which we were first able to apply our measures of drinking, and are also the years in which drinking usually begins, we are encouraged to find that at least in the most critical of our drinking measures (problem drinking), there is some evidence of HLAY curriculum impact over time.

As before, we will offer our readers a few visual illustrations of some of the comparisons that lie behind our findings in Table 8. These illustrations are found on Graphs 6 through 11. Differential "treatment" or exposure patterns are indicated by the X and O symbols: XXX refers to those students exposed to the HLAY curriculum and tested all three years ("pure experimentals"); OOO refers to those tested all three years but never exposed ("pure controls"); and the mixed conditions (XOX or OXX) refer to those exposed during the third year and one of the other two years, but tested in all three years. The XXX condition occurred only in the youngest cohort (4-5-6), due to variations in curriculum requirements in the older (junior high) grades. In Graph 6, for example, we can see what curriculum impact occurred all three years with respect to the knowledge dimension, with the top half of the graph referring to the youngest cohort (4-5-6) and the bottom half referring to the oldest (6-7-8). Both halves of Graph 6 come close to the "ideal" pattern that one would expect from repeated exposure to the HLAY program: In the various experimental conditions (XXX, XOX, or OXX), curriculum impact tends to increase somewhat each year (or at least hold steady), always reaching the maximum in year three. Meanwhile, in the control condition (OOO), even if there is some increase, it does not keep pace with that in the experimental conditions. Such an "ideal" pattern implies some incremental (or "cumulative") curriculum impact each year beyond what has occurred in previous years.

In Graph 7, on the self-esteem dimension for the 6-7-8 cohort, we see an interesting variation on this "ideal" pattern: While the test scores in the control condition remain virtually unchanged over the three years, both of the two experimental conditions (XOX and OXX) show increases. They start out in year one so close to the controls that the differences among the three conditions are not statistically significant. However, the two experimental groups rise steadily in self-esteem scores until their differences from the control group reach statistical significance in the third year. This is again a reassuring instance not only of curriculum impact year to year, but of increasing impact with additional exposure. Graph 8, which deals with the ability to select responsible alternatives in the decision-making process, shows a different but also important pattern to consider in any evaluation of curriculum impact: In years one and two, the three treatment (exposure) conditions remain so close together that statistical significance is not reached, even though the control condition remains lowest in aggregate test scores all three years. By the third year, however, statistical

significance has been reached in the differences across exposure conditions, but not because of increases in the two experimental groups. Indeed, all three groups have actually experienced systematic declines across the three years; but even with the declines, the experimental groups stayed ahead of the control group in each year. This pattern suggests that while peer and other influences in the early teen years may be counteracting the impact of the HLAY curriculum, those influences do seem to be blunted somewhat by curriculum exposure. In other words, while the youngsters during those troublesome years may be inclined anyway to lose some of the conforming and compliant tendencies they had as children, they will not lose as much if they are exposed to the HLAY program.

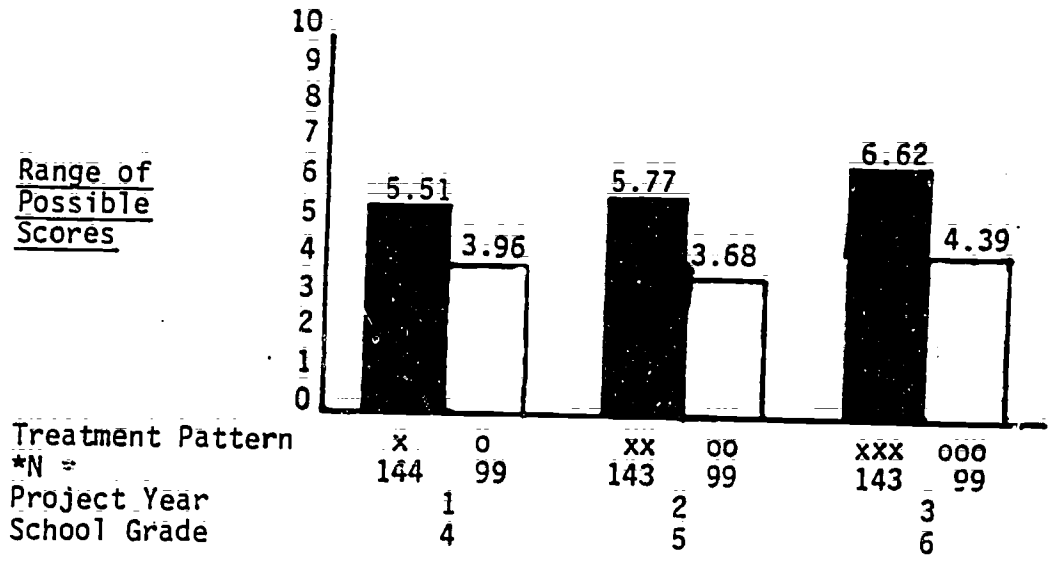
This same general tendency to lose ground in the 6-7-8 cohort can be seen in the bottom half of Graph 9, which deals with the ability to generate alternatives in the decision-making process. While the pattern is somewhat erratic this time, especially in the second year, there is a definite drop in scores for the control group (OOO) from the first to the third year, while the drop is far less serious for the one experimental condition (XOX), and both experimental conditions in year three seem to be on the rise from year two. In the top half of the same graph, for the 4-5-6 cohort, we see increases across the years in this skill for both control and experimental groups, but the scores increase more for the experimental group. We think that these last two graphs (No. 8 and No. 9) point to the interesting observation that the HLAY curriculum can be expected to have its impact in somewhat different ways, depending upon the age groups involved, and upon the particular point in age at which youngsters are exposed to it. The so-called "tempestuous teens," which start in the junior high years, can be expected to bring for nearly all youngsters a time of boundary-testing, some rebelliousness, and a general tendency, reinforced by the peer culture, to reject adult ways of thinking. In this period of life, perhaps the most that can be expected of any school-based prevention program is to slow down such tendencies, or to neutralize them somewhat. It is likely also that the earlier the prevention effort begins, the greater will be its impact on the teen years. In that connection, one can see in these graphs that with very few exceptions the XOX exposure condition (exposed the first year but not the second) yields relatively stronger curriculum impact than the OXX condition (exposed for the first time only in the second year).

In Graph 10 we have illustrated our findings about HLAY curriculum impact on attitudes for the 4-5-6 cohort, even though the evidence for impact here did not rate a plus on Table 8. In this graph, we are dealing only with the total score summing up curriculum impact across all the specific attitudinal areas listed on Table 8. We have found repeatedly in this project that attitudes relating to alcohol use (like most other attitudes) are so hard to change, and/or that small changes are so hard to measure, that the total score is often the best indicator. The situation in Graph 10 is further complicated by the fact that our instruments used more elaborate attitude measures starting in grade 6, so that the measurement basis in year three for the 4-5-6 cohort was different from that in years one and two. That may well be the reason, in fact, that the control/experimental difference for this cohort did not reach statistical significance in the third year, and thus did not rate a plus on Table 8, as it would have on the basis of the first two years. As Graph 10 shows, the gap between the experimentals and the controls widens from the first to the second project year. In the third year, the results still favor the experimentals, but not by as much with the more complex battery of attitude measures.

GRAPH 6

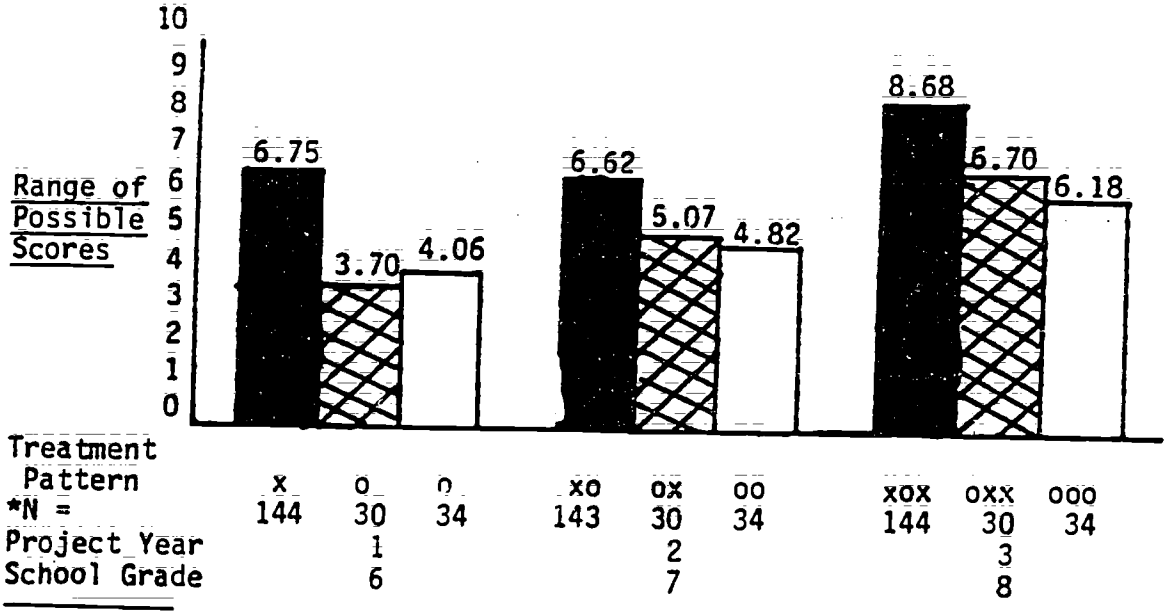
CUMULATIVE PROJECT IMPACT

A. Grade Cohort 4-5-6: KNOWLEDGE



CUMULATIVE PROJECT IMPACT

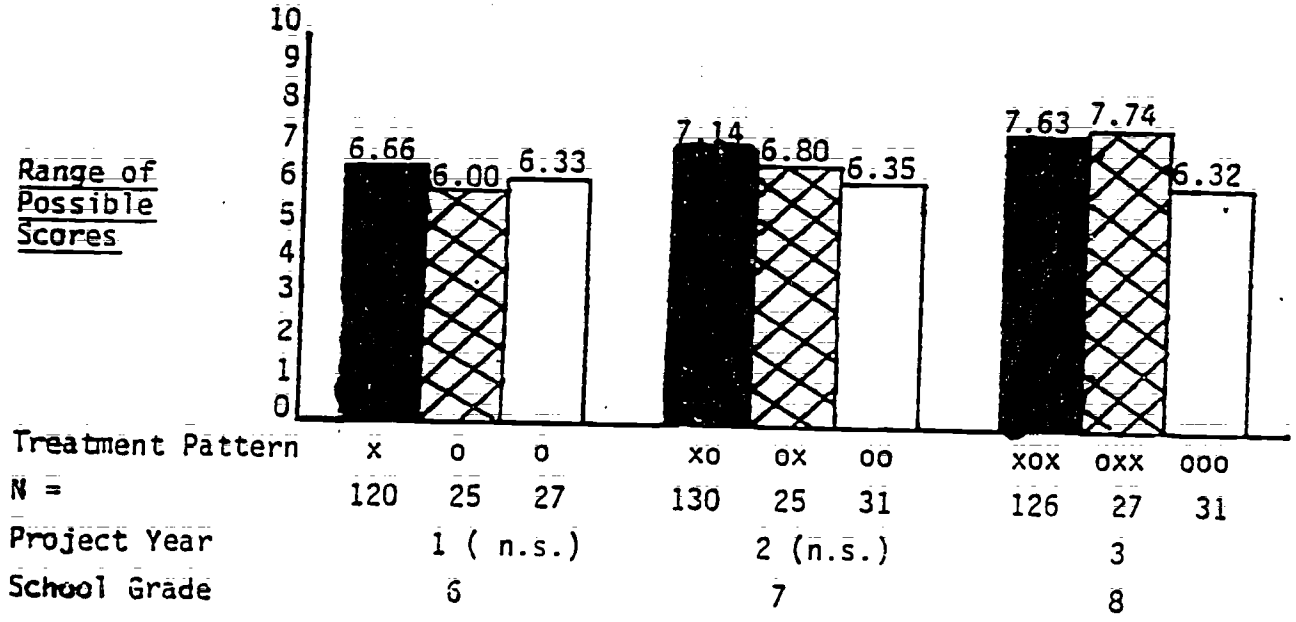
B. Grade Cohort 6-7-8: KNOWLEDGE



GRAPH 7

CUMULATION PROJECT IMPACT

Grade Cohort 6-7-8: SELF-ESTEEM

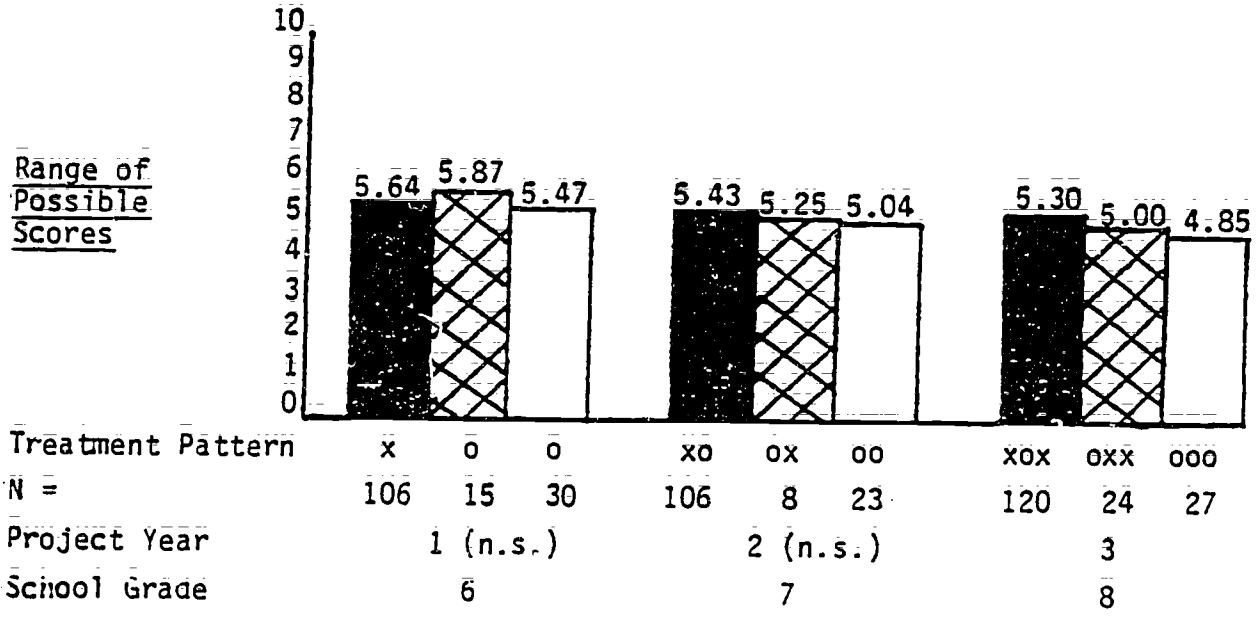


NOTE: n.s. = not statistically significant

GRAPH 8

CUMULATION PROJECT IMPACT

Grade Cohort 6-7-8: SELECTING RESPONSIBLE ALTERNATIVES

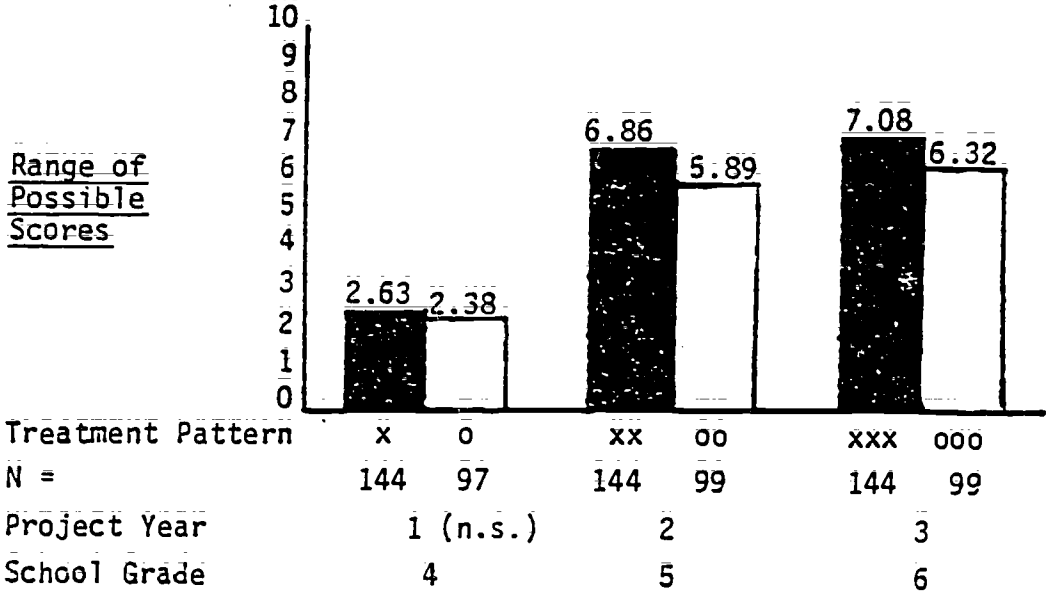


NOTE: n.s. = not statistically significant

GRAPH 9

CUMULATIVE PROJECT IMPACT

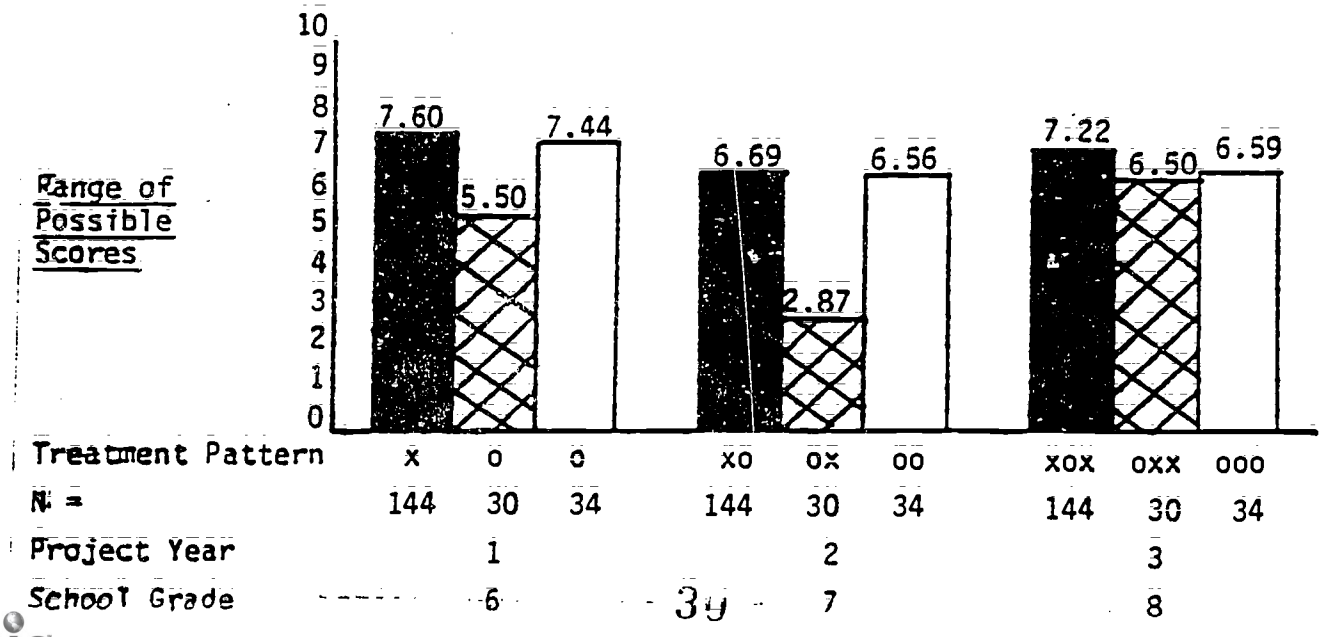
A. Grade Cohort 4-5-6: GENERATING ALTERNATIVES



NOTE: n.s. = not statistically significant

CUMULATIVE PROJECT IMPACT

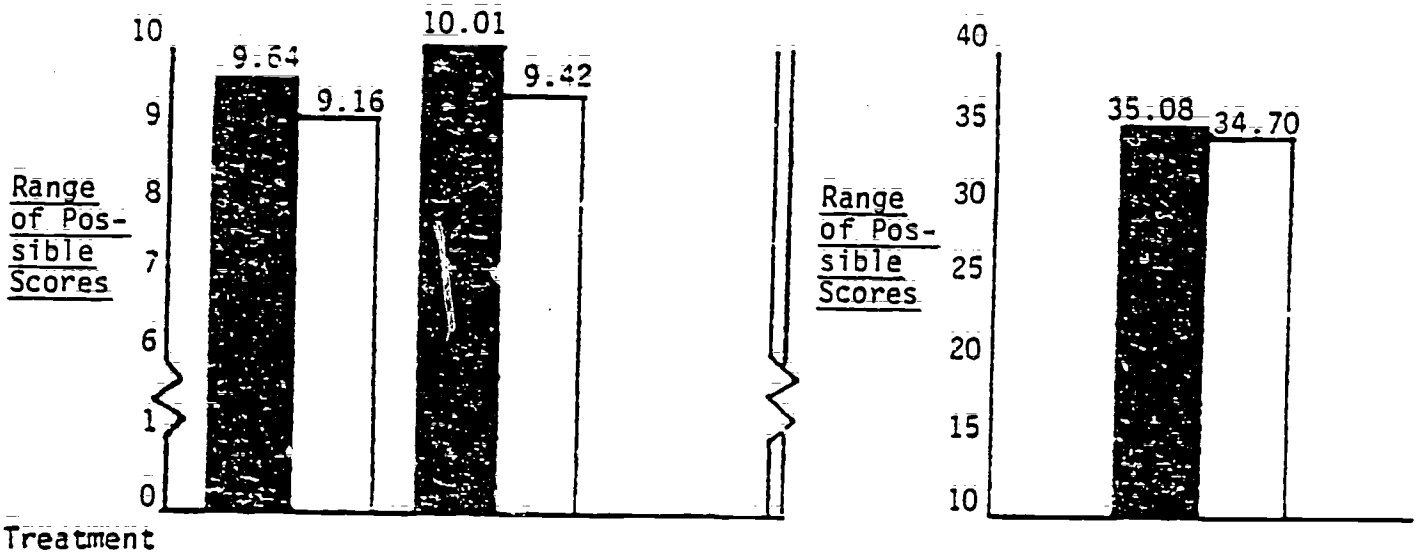
B. Grade Cohort 6-7-8: GENERATING ALTERNATIVES



GRAPH 10

CUMULATIVE PROJECT IMPACT

Grade Cohort 4-5-6: ATTITUDES (Total Score)



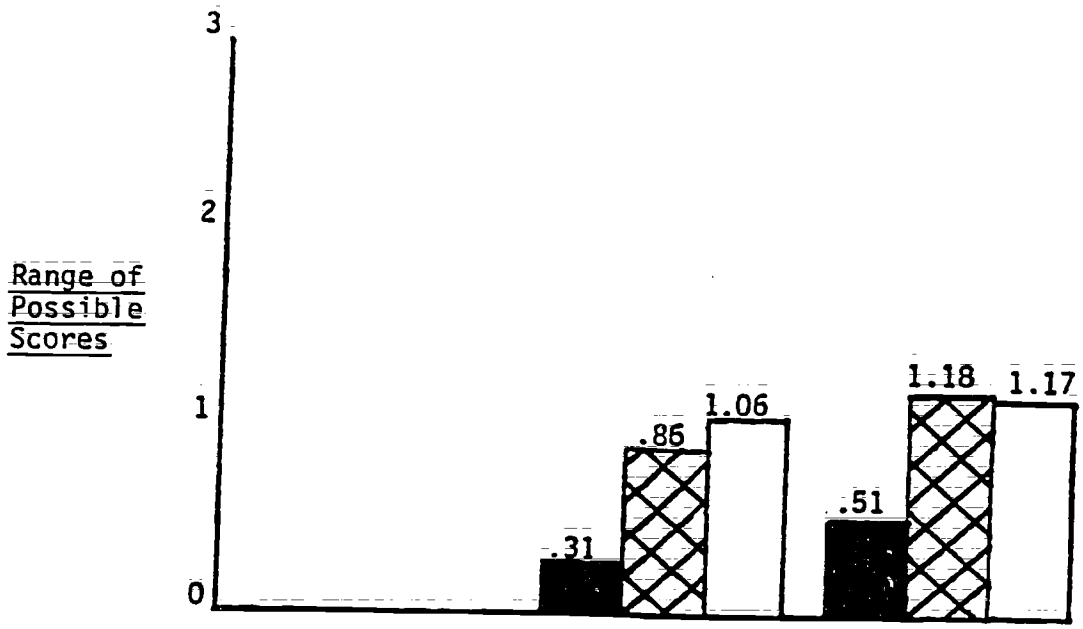
Treatment Pattern	x	o	xx	oo	xxx	ooo
N =	137	95	142	96	133	86
Project Year	1		2		3 (n.s.)	
School Grade	4		5		6	

NOTE: n.s. = not statistical significant

GRAPH 11

CUMULATIVE PROJECT IMPACT

Grade Cohort 6-7-8: PROBLEM DRINKING



Treatment Pattern	(no measure)	xo	ox	oo	xox	oxx	ooo
N =		140	22	31	138	28	29
Project Year	1		2			3	
School Grade:	6		7			8	

NOTE: In this graph, the smaller numbers indicate fewer problems related to alcohol use.

Finally, in Graph 11, we can see some limited evidence for HLAY curriculum impact on actual drinking behavior in the 6-7-8 cohort. Remember that this is the first or youngest cohort for which we have more than one year's measures for different aspects of drinking behavior (our test instruments introduced these starting only in grade 7, the grade or age at which significant drinking first starts for most youngsters). Note this time that the smaller numbers are the more favorable ones. That is, the lower decimal figures indicate lower levels of problem drinking. Here again, in this graph, we can see the kind of "built in" tendency to "lose ground" with youngsters as they move into the junior high years. Problem drinking of various kinds is admitted by the youngsters in all three of the curriculum exposure conditions. However, those in the XOX experimental group increase in their problem drinking levels much less than those in the totally unexposed control group (OOO). The other exposure pattern (OXX) doesn't make out so well. While still below the control group at the start, it catches up to virtual parity in the third year. This difference between the two experimental conditions indicates again that early curriculum exposure, before the junior high years (XOX), produces a more favorable prognosis than later exposure (OXX), even if exposure takes place two years out of three in both cases. The differential between the early-exposed (XOX) and the never-exposed (OOO) in the two years shown in Graph 11 is quite substantial.

All in all, the long-term impact of the HLAY upon cohorts of youngsters exposed to it over a 3-year period is encouraging at several points, as indicated both by Table 8 and by the illustrative graphs that we have just reviewed. The oldest of the three cohorts analyzed here (6-7-8), which spanned the crucial years of the entry into adolescence, showed especially promising results in some of the affective and behavioral dimensions of program outcome. On the other hand, it must be admitted that evidence for curriculum impact was not in any sense systematic or across the board; it was the exception rather than the rule, when we consider all of the grade levels and all of the various dimensions of the curriculum. Also, the Ns for some of the control and mixed conditions in the older cohort are quite small, although that situation does not impair the credibility of our conclusions nearly as much with longitudinal as it would with cross-sectional analysis. Finally, our readers should understand clearly that what we have presented here regarding long-term curriculum impact is not conclusive, for much additional analysis remains to be done by means of truly longitudinal designs across the various panels of our data. Plans for these more complicated analyses, and the additional information we would expect them to yield, are addressed to some extent in the Scientific Appendix to this report.

MAJOR FACTORS AFFECTING CURRICULUM IMPACT

Determining whether or not a given curriculum or other prevention program will "work" is not really so simple a matter as comparing the test results of students who are or are not "exposed" to it. Many factors intervene between the initial design of a program and its "outcome," however that is measured. One important intervening variable is whether or not the program is itself soundly conceived: Do the elements it contains really have anything to do with the problem it is trying to ameliorate? If so, how much? Applied to the HLAY program here under scrutiny, such questions would take these forms: Are there empirical grounds for believing that knowledge and information about alcohol, self-esteem, coping and decision-making skills, and certain attitudes have anything to do with drinking behavior? If so, how much when compared to the other factors operating in young people's lives, such as parent and peer influences, demographic traits, and so on?

A second obvious question that has to be considered in assessing how well a program "works" is how well it was "delivered." In the present instance, that translates into a number of questions about the teachers, who were the chief agents or means of delivery: How were the teachers trained? How much good did their training do them? How much commitment to the HLAY program did they bring to their classrooms? How conscientious were they in implementing the program in accordance with the training they had received? What difference, if any, did variations in teacher characteristics make for how well the students did on the evaluation tests or instruments? While we still have much analysis to do on all such questions about teachers, we do have some relevant information to pass along at this point. We will first, however, address the questions in the above paragraph about the relevance of the various dimensions of the HLAY curriculum itself to drinking behavior; or, put another way, how much "potential" the curriculum itself contained for making any impact on drinking.

The Potential for Curriculum Impact on Drinking Behavior

As will be clear from our discussion above, and from Table 1, our data-collection instruments contained measures for a variety of factors influencing the youngsters in addition to the HLAY curriculum. These other factors included the so-called demographic variables like age, sex, and race, as well as the major socializing variables like parents, peers, and religion. All these other factors exert influence on a child's life before, during, and after his/her experience with any school program. Therefore, in evaluating such a program, it is important to determine how much "room" there is for the program to have an influence, given the power of the social and demographic factors, already at work, which have "pre-empted," as it were, much of the influence that will determine how a young person deals with alcohol. One way of estimating the relative potential impact of the HLAY curriculum, compared to all these other influences, is to throw them all into a regression equation, with various kinds of drinking behavior measures as the dependent or "outcome" variables. Table 9 shows the results for this method of estimating the potential for curriculum impact, based upon our cross-sectional subsamples. (The subsamples varied in size somewhat because of occasional missing data or non-response to specific measures.)

The decimal figures in Table 9 may be interpreted simply as the net proportions or amounts of the variation in the different measures of drinking behavior that can be accounted for by knowledge and information about alcohol, self-esteem, decision-making skills, and attitudes about alcohol, after removing the effects of social and demographic variables. The so-called "curriculum variables," whose effects are shown in this table, are the cognitive and affective traits of the students that the HLAY program is aimed at influencing. Quite aside here from the question of whether or not that program really does influence those student traits, we are addressing the question of how much change in the (aggregate) drinking behavior of students could the HLAY program make, even if its impact was at maximum levels. As the figures on this table indicate, this potential for curriculum impact ranges from a low of .04 (4%), for Grade 7 monthly quantity and frequency of drinking, to a high of .26 (26%), for Grades 10-11 annual quantity and frequency. Generally speaking, the figures are modest, indicating that when influences of family, peers, religion, age, sex, race, and many other things are taken into account, there is a limited amount of room left for any alcohol education or prevention program to have much effect. The figures are smallest for the lower and junior high grades, then they reach a peak, in general, in the central

TABLE 9

INDEPENDENT CONTRIBUTION OF TOTAL SET OF CURRICULUM VARIABLES TO DRINKING BEHAVIOR OF STUDENTS NOT PREVIOUSLY TESTED OR EXPOSED TO CURRICULUM

	Gr 4	GR 5	GR 6	GR 7	GR 8	10-11	Gr 12
Range of N's for these Analyses							
Maximum	685	239	244	209	541	118	251
Minimum	685	239	244	198	480	84	193
Childhood Drinking Behavior	.05	.14	.10				
Adolescent Drinking Behavior							
Irresponsible Uses				.12	.08	.12	.12
Problem Drinking				.07	.05	.11	.14
Current Drinking Situation				.08	.06	.12	.10
Expectations about Drinking				.09	.10	.12	.11
Monthly Frequency				.13	.09	.24	.17
Yearly Frequency				.12	.09	.21	.09
Quantity				.07	.05	.19	.09
Quantity-Frequency Monthly				.04	.06	.23	.19
Quantity-Frequency Yearly				.05	.06	.26	.18

Note: The data for this table come from the unexposed or "naive" students described on page 8 and Table 3. The severe attrition in Ns from that table to this table are accounted for by the "list-wise deletion" procedure built into our regression program, which has the effect of eliminating from the analysis any and all students who did not respond to all of the questionnaire items. The variation in Ns from maximum to minimum in the above analyses is due to still further attrition from any subject's failure to respond to all items constituting a composite variable; consequently, the effective Ns are smaller for the multi-item variables above than for those measured by a single item. The R^2 values were obtained for the set of knowledge, self-esteem, decision-making, and attitude variables after partialling out all other variables (demographic and non-curriculum) measured in this project. Grade 9 is omitted from the table because the Ns were rather small (this preliminary analysis essentially requires that each student included in the analysis respond to all items on the questionnaire).

high school years (10-11), followed by a decline again during the senior year of high school. The relatively high potential for curriculum impact on drinking behavior in grades 10 and 11 is somewhat ironic and frustrating in view of our finding (see Table 7) that actual curriculum impact on the cognitive and affective traits of students is more likely to occur in the earlier grades (especially attitudes and self-esteem). In examining Table 9, however, it is important to keep in mind that the figures are really quite parsimonious estimates of the room for potential curriculum impact. They are the net residual figures generated by regression equations from which other plausible influences upon youthful drinking behavior (social and demographic) have already been eliminated. In other words, these are estimates of what is left for the curriculum to impact, after a variety of other influences on the youngsters' lives have been taken into account. From this point of view, though the figures do not often reach even 10% (.10 or more), they are by no means negligible and should be understood as generally supporting the theoretical basis on which the HLAY curriculum is built.

Graphs can provide somewhat more visual illustrations of the relative potential we have been discussing. In Graph 12, we can see that the factors addressed by the HLAY curriculum (knowledge, self-esteem, decision-making, attitudes) could affect 7% of the problem-drinking in grade 7, out of the 35% that the total equation could account for. At grade 8, the corresponding figure is 5% out of 16%; at grades 10-11, 11% out of 35%; and at grade 12, 14% out of 38%. In the last two graphs (13 and 14), quantity and frequency of drinking on a monthly and an annual basis, respectively, we see the same general pattern with respect to the relative potential of the curriculum-related influences. However, in these two graphs we can see a striking increase, after junior high school, in the amount of variation in drinking patterns that can be explained merely by the combined social, demographic, and curriculum-related factors (about 50% or more). Also, the cognitive and affective factors addressed by the HLAY curriculum now account for much more of the variation, and therefore of the potential for program impact.

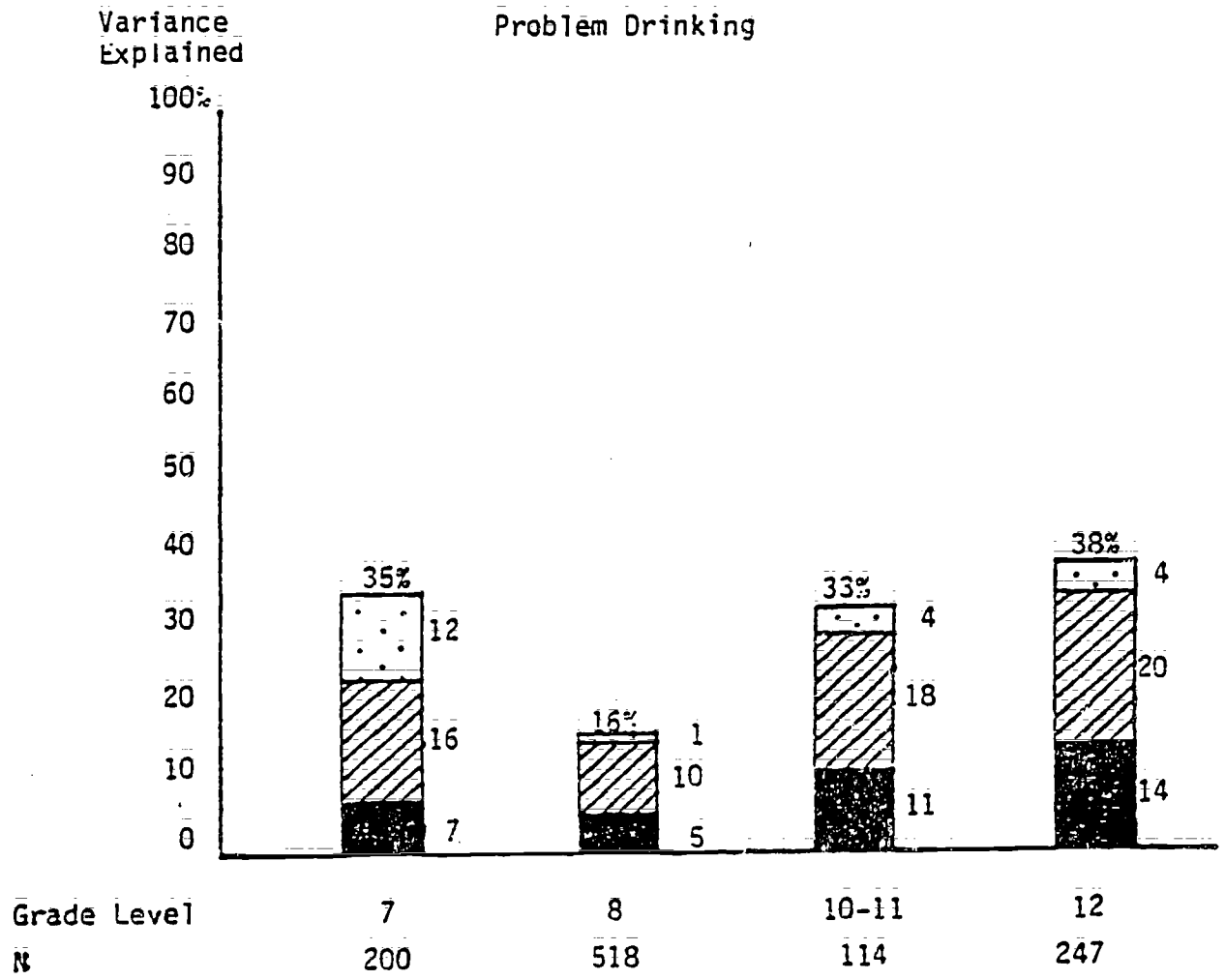
Teachers and Curriculum Implementation

Altogether, across the three project years, 430 different teachers were involved in providing our data, of whom about 3/4 represented "experimental" classrooms and the rest "control" classrooms. Our various evaluations of the teacher-training component of the HLAY program have indicated that in general the training has a definite impact of the desired kind in getting teachers ready to employ the HLAY materials and philosophy in their classrooms (see Rankin, et al., 1978, and Tarnai, et al., 1981). At the same time, however, teacher attitudes toward various aspects of alcohol use did not yield to training as much as most other teacher traits did. Furthermore, the outcomes of teacher training were affected by certain pre-existing factors, such as the amount of classroom experience a teacher had had and the level of enthusiasm which the teacher brought to the training workshops. We should not be surprised that training affects different teachers differently, or that teacher attitudes toward training and toward alcohol should account for many of the differences in training outcome. Nor, in view of all that, should we be surprised to find, as we did, that in their actual employment of the HLAY curriculum in the classroom, teachers had less overall impact on student attitudes than upon any other single dimension of the curriculum. Indeed, in the follow-up questionnaires that we administered, teachers tended to rate the attitude-oriented parts of the HLAY curriculum as among the least important parts. Social scientists know that attitudes are

GRAPH 12

COMPARATIVE RELATIONSHIP OF CURRICULUM VARIABLES AND OTHER VARIABLES WITH DRINKING BEHAVIOR

(Figures are Percents of Variance Explained in Drinking Behavior)



= Curriculum Variables

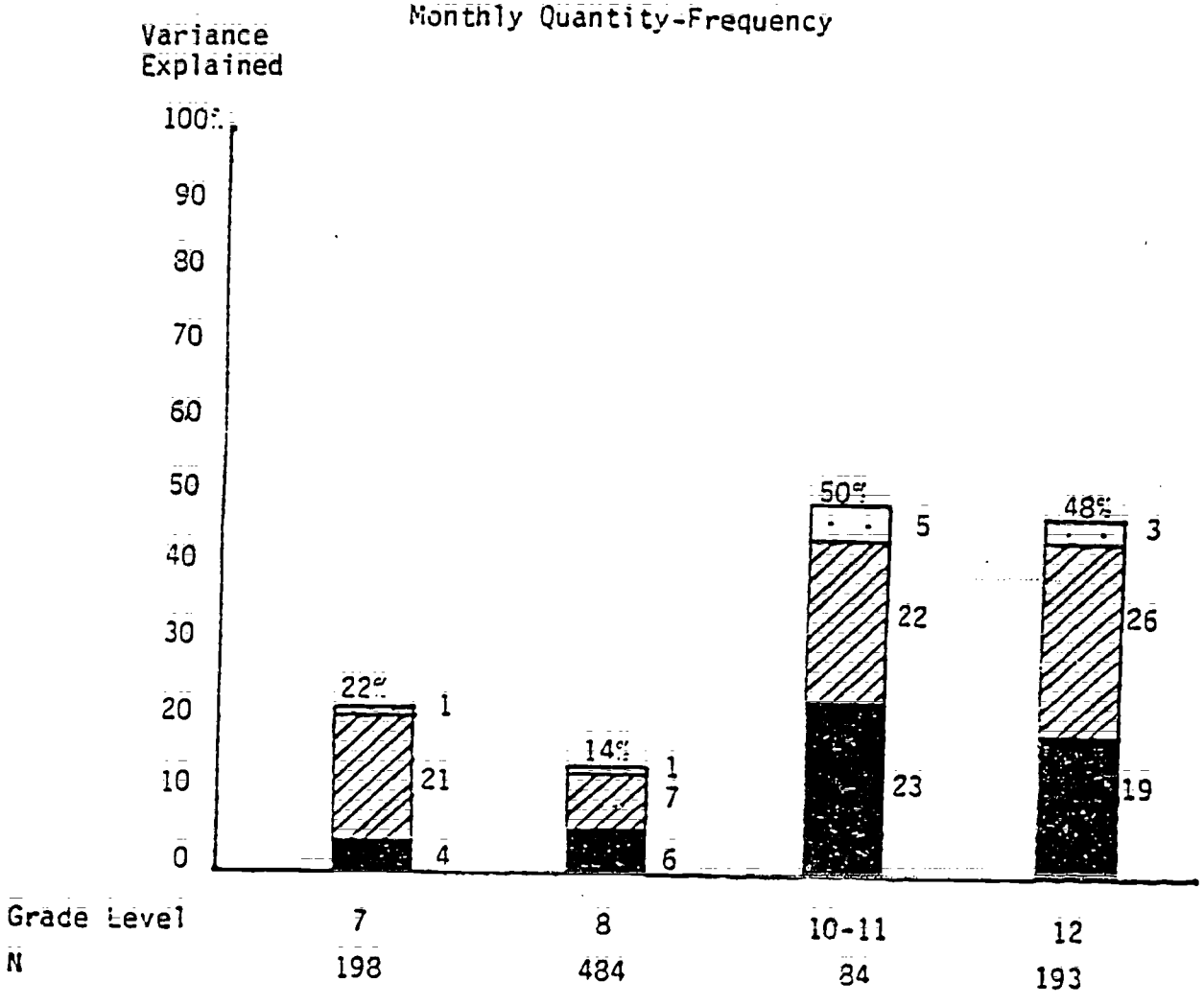
= Social Variables

= Demographic Variables

GRAPH 13

COMPARATIVE RELATIONSHIP OF CURRICULUM VARIABLES AND OTHER VARIABLES WITH DRINKING BEHAVIOR

(Figures are Percents of Variance Explained in Drinking Behavior)



= Curriculum Variables

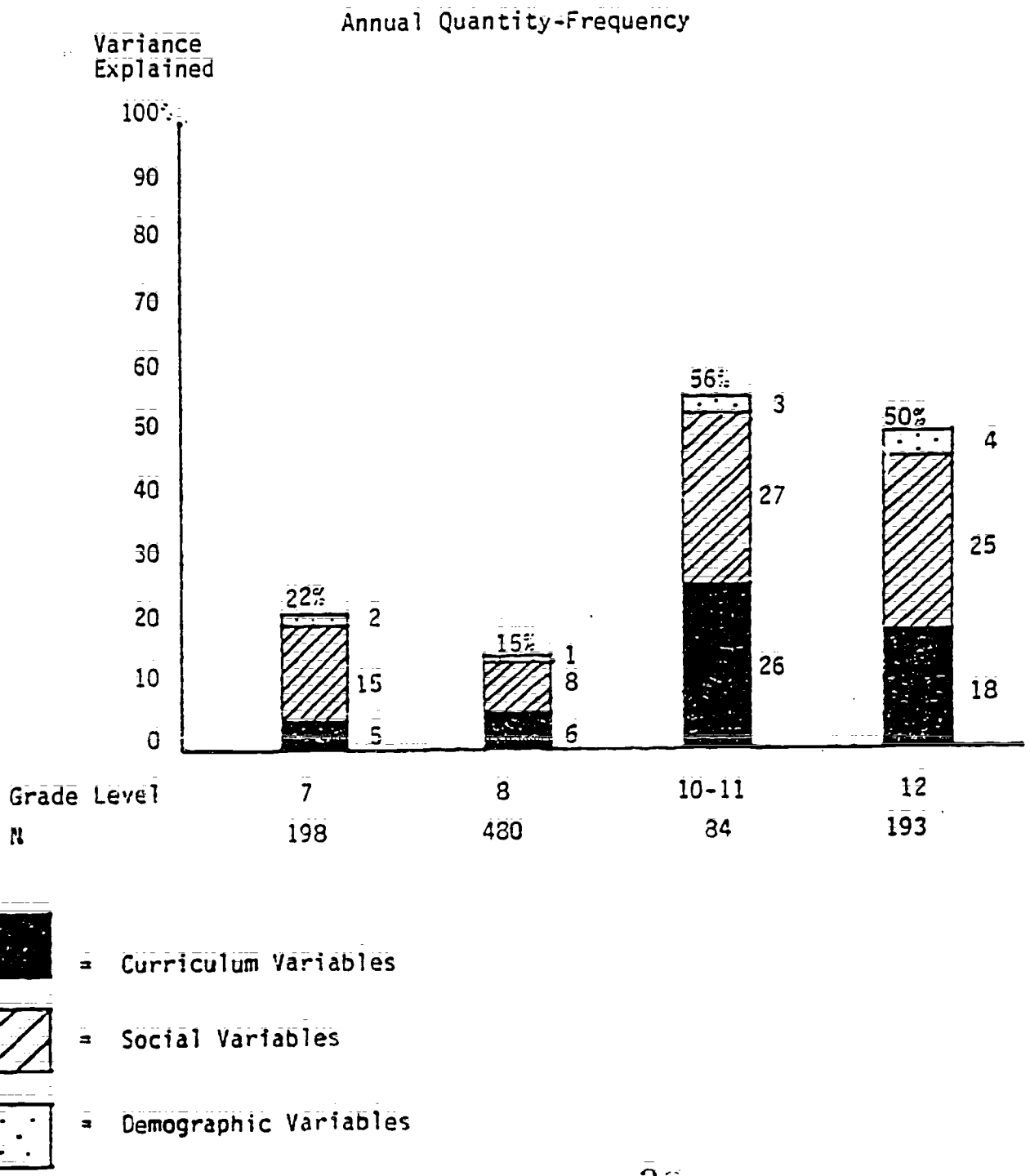
= Social Variables

= Demographic Variables

GRAPH 14

COMPARATIVE RELATIONSHIP OF CURRICULUM VARIABLES AND OTHER VARIABLES WITH DRINKING BEHAVIOR

(Figures are Percents of Variance Explained in Drinking Behavior)



hard to change, especially for any significant period of time, and even more so to the extent that the change agents (teachers, in this case) are themselves ambivalent.

With respect to the actual implementation of the HLAY curriculum in the classrooms, there were at least three fairly mechanical factors we attempted to measure that can usefully be discussed here: (1) The number of classroom periods of curriculum exposure (out of the 15 called for) that teachers actually conducted; (2) The relative amounts of emphasis actually given to the various dimensions of the curriculum: knowledge, self-esteem, decision-making or coping, and attitudes; and (3) The amount of time that was permitted to elapse between completion of the curriculum and the administration of the test instruments. Our data for these factors came from brief reports that teachers filled out and returned with each batch of student tests. We will address each of these factors first in a brief descriptive way.

Total time of curriculum exposure: There was enormous variation across the classrooms in our project in the total number of classroom periods devoted to the HLAY project. Few teachers achieved the stipulated goal of at least 15 periods, but there was a definite tendency toward more periods of exposure at some grade levels than at others: If we count at least 14 periods as "virtual compliance" with the HLAY goal, then we can say that project-wide half of the teachers met that goal at grades 4, 7, and 9, while at grades 10 and 11 that goal was met by two-thirds of the teachers. At the other grades we studied, compliance with the goal of 15 (or 14) periods of curriculum exposure was reached by a third of the teachers or fewer; indeed, at the 12th grade level almost no teachers spent more than 10 periods.

Differential emphasis on various dimensions of the curriculum: Each time the teachers taught with the HLAY curriculum in their classrooms, we asked that they indicate on a questionnaire how much emphasis they gave (respectively) to the knowledge, self-esteem, decision-making, and attitudes dimensions of the curriculum. The response categories for each dimension were "much," "some," "little," or "none," admittedly a very rough measurement system. As one might expect (and we expected), the extremes ("much" or "none") were more readily interpretable categories than were the intermediate ones. However, we found that almost no teachers anywhere would admit to having given any of the curriculum dimensions no emphasis. In the case of the knowledge dimension, the most common response was "much" emphasis, reaching close to 100% of the teachers, indeed, except in the lower grades (4 and 5). That is not surprising, since this aspect of the curriculum is not only the most conventional part of alcohol education programs, but is also the easiest to teach. For the other three dimensions, the modal response category was "some" emphasis, though for self-esteem there was an approximately even split in grades 4 and 7, where about 50% of the teachers were found each in the "much" and the "some" categories of emphasis. This same kind of even split occurred in grades 10 and 11 both for decision-making and for attitudes. Otherwise, almost all teachers in all grades and for all dimensions claimed at least "some" emphasis.

Interval between curriculum exposure and testing: In the expectation that the performance of the students on our measuring instruments (tests) might be affected by any appreciable delay in the administration of the tests, we tried to keep track of this factor as a variable to consider in our evaluation. The main empirical "break" in our data on this subject came after a delay of three days: that is, the majority of our teachers got their tests administered within three

days or fewer of the completion of the HLAY curriculum. At some grade levels, however, there were appreciable numbers of teachers (and thus of classrooms) for which the lapse was considerably longer than three days. The usual reasons for longer lapses were: (1) A vacation break of some kind intervened before testing; (2) the HLAY curriculum materials and exercises were integrated (and therefore scattered) across a more general course or unit (like health education), the overall testing for which was done at the end of the unit; (3) Logistical problems intervened, such as failures of the clerical staffs and/or of the postal service to get the necessary tests delivered to the teachers on time; or other miscellaneous difficulties. For whatever reasons, however, half or more of the "experimental" teachers failed to administer tests within three days in grades 4, 5, and 12. Otherwise, the three day interval (or less) was the rule in about 80% of the experimental classrooms at all the other grade levels.

Some implications of differentials in teacher performance for student performance: While we have yet to complete a thorough and comprehensive multivariate analysis of the part played by teacher-related variables in curriculum impact on students, we do have some grounds for suspecting that teachers can make a difference. Just how much difference, however, and under what circumstances, is not yet clear. Many teachers at all grade levels failed in one way or another to implement the HLAY curriculum according to design, and this could be one important reason for the modest impact that the curriculum has had in most respects. On the other hand, the statistical analyses we have carried out so far do not show much systematic difference in curriculum impact according to teacher performance.

Some additional evidence bearing upon the relative importance of teacher effort and commitment in the implementation of the HLAY curriculum is to be found in the special sub-study we did on that matter during the third year of our project. During the first two years of the project, we had tried hard both to monitor teacher behavior in the classroom and to enhance teacher compliance with the particulars of the implementation design. In doing so, we had had to rely primarily on our ongoing communications with the teachers by mail and by telephone, except for the formal instruction they had been given at the outset and in the subsequent "booster" sessions of training at the start of each new school year. Our long distance from the sites where the HLAY program was being implemented made it logistically and financially prohibitive for any members of our research staff to make observational visits to the project classrooms and monitor the process of delivery or implementation first hand. However, as it became obvious that teacher compliance with the implementation design for the HLAY curriculum would inevitably be quite variable across the project, and would seldom be complete, we began to consider whether any more could feasibly be done to gain greater teacher commitment and compliance, so that the HLAY curriculum could more often be implemented with the requisite number of classroom exposures, the requisite emphasis on each curriculum component, and the appropriate teaching methods employed, as called for in the HLAY design. We questioned too how much difference such an enhancement of teacher commitment and compliance would make in curriculum impact upon the students.

At the suggestion of our colleagues at the Educational Service District No. 121, who had developed the HLAY program and trained the teachers, we selected a few teachers that they recommended to us as more skilled and committed to the program than most of the teachers. Special contracts were concluded with these selected teachers, in which we agreed to pay them $100. each for the strictest possible compliance with the implementation design of the HLAY curriculum, and they agreed further to submit to a number of monitoring visits to their classrooms by professionals from the ESD staff. Altogether, three of these selected teachers,

representing four different grade levels, entered into these contracts and generated enough data for analysis. We analyzed the data from their classrooms separately and made comparisons of their data with project-wide norms for evidence of differential curriculum impact on students in the immediate or short-term sense.

To speak in generalities, we would have to say that the results of this comparison were mixed, as we have usually found the evidence on teacher impact to be. At the 4th grade level, there were no significant differences for students of the selected teachers, either on a pre-test/post-test basis, or as compared with project norms for that grade, with the exception of some modest improvements over project norms in the area of decision-making skills. In grade 5, the picture was but little better: test results showed significant gains by the students of the selected teacher on a pre-test/post-test basis in a couple of the attitudinal areas and in one of the decision-making skills; but even the post-test results did not compare favorably with project-wide norms for fifth-graders. In grade 6, the students of the selected teacher in general showed no improvement, either on a pre-test/post-test basis, or by comparison with project-wide norms for that grade, with the exception of one of the decision-making skills, where the post-test scores were clearly superior both to pre-test and to project norms. Finally, in the 10th grade, the only secondary level at which we had any special or selected teachers, the students showed no evidence of having benefitted by any special teacher effort, either on a pre-test/post-test basis, or by comparison with project norms for that grade. Such evidence as there was, then, for the importance of strict teacher compliance with the main requisites of curriculum implementation, was more noteworthy in the elementary grades than at the secondary level.

CONCLUSION

This report has covered the most general findings of our evaluation of the "Here's Looking at You" Alcohol Education Program during the period 1978 - 1981. An earlier report, dated December, 1973, covered our evaluation of a field-testing phase, during which various formats of teacher training were assessed, measures were developed for evaluating program impact on students, and limited evaluations were actually carried out of student impact, on an immediate or one-time basis, in various grade levels and locations. This present report goes beyond the earlier one by covering evaluation research that was based on (a) an elaborate quasi-experimental design; (b) large quantities of student data; (c) a longitudinal component, in which some of the students were followed individually and measured repeatedly over a 3-year period; (d) consideration of important factors and influences from outside the HLAY program itself, such as parents, peers and demographic factors; and (e) some assessment of the HLAY program impact on actual student drinking behavior, not heretofore possible without the longitudinal data.

We assessed the impact of the HLAY alcohol education curriculum on students, both on an immediate or short-run basis and over a 3-year period. In an immediate sense, after exposure to the HLAY unit only one time, curriculum impact on students was definitely found for all students in the area of knowledge or information about alcohol and alcoholism. In addition, we found that self-esteem was favorably affected by short-term curriculum exposure in grades 5, 6, and 7, and that decision-making skills of various kinds were improved for youngsters, starting in grade 6 and on into junior high and high school. As we had always found in our preliminary studies, attitudes relating to alcohol use were least affected by exposure to the HLAY unit, but there was evidence in our research that students from grade 8 on down were changed somewhat toward an attitudinal stance favoring

moderate drinking (as opposed to abusive or excessive drinking). Short-term curriculum impact on actual drinking behavior could not logically be evaluated, since most of the behavior reported would have occurred prior to exposure to the HLAY unit. However, we did find reason to believe that high levels of pre-existing drinking behavior negatively affected student receptivity to the HLAY curriculum and its objectives. These findings on short-term curriculum impact were more promising than any that had emerged from our earlier and more preliminary studies, and the findings did not differ appreciably from one school district to another.

Our research on the continuing and/or cumulative impact of the "Here's Looking at You" program, over a 3-year period, produced less encouraging and less decisive results. Again, as always, the curriculum produced a clear and growing impact on the knowledge and information students retained about alcohol and alcoholism across time. Much less general, but still observable, were improvements sustained across time in some of the decision-making skills. The most encouraging results were found for the age cohort 6-7-8, that being the longitudinal subsample first measured while they were in the 6th grade and then during the succeeding two years. In this cohort, we found convincing evidence of continuing (and often even increasing) HLAY curriculum impact across the three years, not only in knowledge, but also in self-esteem and in some decision-making skills. Where our measures of drinking behavior were concerned, we found definite and strong curriculum impact in this cohort across time upon problem-drinking, but not otherwise upon the quantity or frequency of alcohol use. The 6-7-8 grade cohort is an especially important one, of course, since it spans the crucial period of entry into puberty and the teenage years. Our data in general have indicated that the drinking attitudes and behavior acquired by the time of entry into high school are likely to endure for some time, so this pre-high school cohort is a particularly important one in which to have found curriculum impact on behavior and on some of the other attributes. This finding also argues for early intervention with school-based prevention programs, well before junior high school. Not that such intervention, even on a continuing or cumulative basis, will prevail over peer and other influences during these crucial years. Indeed, we observed in this 6-7-8 cohort that in some important respects (notably problem-drinking behavior and responsible decision-making), school influences actually lost ground compared to other influences; however--and this is very important--students in that cohort repeatedly exposed to the curriculum lost less ground than those not exposed at all, and tended to lose it more slowly. Thus, school-based prevention programs like HLAY may not be able to prevent the increase of certain "natural" tendencies among teenagers toward boundary-testing with respect to alcohol and many other things. However, such prevention programs in the schools may well be able to inhibit or blunt such tendencies.

Having said all that, we are still left with the more general observation that the impact of the "Here's Looking at You" Alcohol Education Program has been apparent only to a limited extent, only in some of its aspects, and primarily in the short term. Impact on attitudes has proved especially rare, or at least difficult to measure. To the extent, however, that impact on drinking behavior itself can be demonstrated, perhaps impact on the intermediate factors like attitudes or decision-making skills is less important. Modest, if definite, impact on drinking behavior does seem to have occurred gradually across time in at least one crucial cohort of pre-high school students. Anecdotal and other unsystematic feedback from the classrooms has indicated also that in general both teachers and students like the HLAY program. In view of such a mixture of findings, we are left essentially with a cost-benefit assessment that will have to be made by school boards and administrators contemplating the adoption of the HLAY program.

In view of the manifest and sustained effort and expense that went into the development of the "Here's Looking at You" program, and into its evaluation, we might well ask why its impact was not more dramatic. Why didn't it "work" better? What could have been done (or could be done in future adoptions) to make it work better? In Table 9, and in the graphs which followed it, we have already seen that the HLAY program was well conceived, in the sense that its major dimensions (knowledge, self-esteem, attitudes, coping, and decision-making) are, at least in the aggregate, clearly related to alcohol use and abuse. At the same time, however, it is just as clear from those same graphs that drinking behavior (alcohol use or abuse) is not nearly so dependent upon these curriculum-related dimensions as it is upon the other influences in the lives of students, particularly their parent, peer, and religious influences. Even if the HLAY program "worked" perfectly, there are stringent limits to the impact it could have on youngsters, in the face of other influences. Still, that potential impact is not negligible, and it tends to increase, even relative to other influences, as the youngsters grow older. There would seem to be an implication here that if the cognitive and affective traits addressed by the HLAY curriculum could be enhanced early, before junior high school, they might loom larger and more important as determinants of drinking behavior later on.

Aside from the potential inherent in the HLAY curriculum, the performance of the teachers in the classrooms is likely the most crucial factor in how fully that potential can be realized. We were satisfied that the teachers in our project had been adequately trained to implement the HLAY in their classrooms according to the design of the developers at ESD-121. Our procedures for monitoring and follow-up of teacher performance were as elaborate and effective as we could make them, given the inherent practical constraints, budgetary, logistical, and diplomatic. Many of our project teachers nevertheless proved unwilling or unable to implement fully the program in their classrooms in accordance with the training they had received and the commitments they had made. As frustrating as this state of affairs may be from an evaluator's standpoint, curriculum developers are likely to be even more frustrated, as they see the program they have so carefully designed being compromised and attenuated by incomplete or slipshod implementation in the classroom. In the face of this frustration, curriculum designers may be inclined to "disown" their own programs, or at least the evaluations of their programs, and to deny the validity of any evaluation that is done without first insuring that the program is being carried out meticulously according to design. In general, this is a sound position to take: there would seem to be no point in evaluating the outcome of a curriculum or any other program that is not being implemented properly.

In this particular instance, however, we have assumed a different position, one that is defensible at least under the circumstances of our particular evaluation project. We understood our charge from the funding agency as one of evaluating a comprehensive school-based alcohol education program under the conditions normally obtaining in a public school system, not under controlled laboratory conditions. To be sure, we wanted to do our evaluation under conditions that would, in every classroom, approach as closely as possible the ideal setting and teacher behavior envisioned by the curriculum designers. However, from the field-testing phase of our work, we knew that there would be many departures from the ideal and much unevenness across the project in teacher performance. Once we had chosen our schools and classrooms, we were "stuck with" whatever teachers went with them. After a period of trying various ways to get

as much teacher compliance as possible with the design for curriculum implementation, we pushed ahead with our plan for outcome evaluation, knowing full well that teacher compliance was going to be problematic throughout the project, or even minimal in many instances. We did so with the following understandings: (1) We would keep track, as much as possible, in as many ways as possible, of variations in teacher compliance, and build that factor into our outcome evaluation as one of the variables to be considered, rather than as a constant to be achieved before evaluation could begin; and (2) We would thereby be faithfully carrying out our charge to evaluate curriculum outcome under "normal conditions" in a school system. In the "real world" in which new school programs have to be tested, nothing is more "normal" than differential teacher commitment and compliance to program design.

To have done otherwise, we feel, would have made little sense in the case of this particular project. One option might have been for us to postpone outcome evaluation until we were satisfied from our process evaluation that we had at least minimum acceptable teacher compliance to design all across our project. Had we done that, we would probably still be tinkering, correcting, and fine-tuning the curriculum implementation process, with very little prospect of being any closer to project-wide compliance, and with neither time nor money left to do any outcome evaluation per se. One must also, then, avoid the danger of tarrying so long on monitoring implementation that one never gets to outcome evaluation. Alternatively, we might have tried to make fundamental changes from time to time in what teachers were doing in their classrooms, or in what we were doing with the teachers, or in some other aspect of the implementation process during the three years of our research, assuming we would have had the power to do so. This would have meant, however, evaluating a constantly changing program, which would have made it difficult to be sure just what we were evaluating--version a, version b, or version n of the program? We had three years of funding to do an outcome evaluation, and the time had come to assess program impact, in the aggregate, as that program would normally be implemented, and with all its unevenness, under typical staff and classroom conditions.

While our evidence so far does not show that variations in teacher performance or commitment "normally" occurring in school settings make much of a difference in curriculum impact overall, our subsequent analyses may yet determine that some student test scores are affected by the sheer number of classroom periods devoted to the HLAY curriculum, or by the length of the delay before testing, or perhaps by other aspects of teacher performance or background. Given the limited potential for this or any school program to compete with other influences in shaping student behavior, therefore, it would in any case behoove any school system which adopts the HLAY alcohol education program to give as much attention as possible to thorough and conscientious teacher implementation of the program in all its aspects, if students and parents are to get at least the program's "best shot."

BIBLIOGRAPHY

Kearney, Kathleen A.
1982 Collecting Longitudinal Data Anonymously: Compensating for Respondent Errors in Self-Generated Identification Codes. Unpublished Ph.D. Dissertation, Washington State University, 1982.

Kearney, K. A., Hopkins, R. H., Weisheit, R. A., and Mauss, A. L.
1982a "Data Loss and Sample Bias Resulting from a Requirement for Written Parental Consent in an Educational Evaluation." Unpublished paper under editorial review.

Kearney, K. A., Hopkins, R. H., Mauss, A. L., and Weisheit, R. A.
1982b "Self-Generated Identification Codes for Anonymous Collection of Longitudinal Questionnaire Data." Unpublished paper under editorial review.

Mauss, A. L., and Hopkins, R. H.
1979 A Manual of Evaluation Guidelines for "Here's Looking at You"--A Model Program in Alcohol Education. Washington State University. May.

Mauss, A. L., Hopkins, R. H., Tarnai, J., and Kearney, K. A.
1980 "Perils and Potshots in Prevention Research: A Response to Staulcup, et al." Journal of Studies on Alcohol 41(9): 959-962.

Rankin, W. L., Mauss, A. L., Tarnai, J., Fagan, N. J., and Hopkins, R. H.
1978 "An Evaluation of Workshops Designed to Prepare Teachers in Alcohol Education." Journal of Alcohol and Drug Education 23(3): 1-13.

Tarnai, J., Mauss, A. L., Hopkins, R. H., and Lawlis, P. M.
1978 Evaluation of the ESD No. 121 Alcohol Education Project of King County, Washington, 1975-1978: Final Report. December, 1978.

Tarnai, J., Magnusson-Fagan, N. J., Hopkins, R. H., Mauss, A. L., and Eichberger, M.
1981 "On Re-Tooling the Teachers: An Evaluation of Teacher Training in Alcohol Education." Journal of Alcohol and Drug Education 27(1).

Weisheit, Ralph A.
1981 Age, Sex, Informal Control, and Adolescent Drinking Behavior. Unpublished Ph.D. Dissertation, Washington State University, 1981.

1982 "The Social Implications of Alcohol and Drug Education: Implications for Program Evaluations." Unpublished paper under editorial review.

Weisheit, R. A., Hopkins, R. H., Kearney, K. A., and Mauss, A. L.
1982a "Substance Abuse, Non-Conformity, and the Ability to Assign Problem Responsibility." Journal of Drug Issues. Spring (in press).

Weisheit, R. A., Mauss, A. L., Hopkins, R. H., and Kearney, K. A.
1982b "Deviance and Responsibility." Unpublished paper under editorial review.