AN EVALUATION OF A GRADE SIX DRUG EDUCATION PROGRAM

by

David H. Erickson

B.A. (Honours), University of Regina, 1979

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in the Department

of

Psychology

C David H. Erickson 1984

SIMON FRASER UNIVERSITY

April 27, 1984

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.

Approval

Name: David Harry Erickson

Degree: Master of Arts

Title of thesis: An Evaluation of a Grade Six Drug Education Program

Examining Committee:

Chairman: Dr. P. Bakan

Bruce K. Alexander, Ph.D.
Senior Supervisor

Ronald M. Roesch, Ph.D.

E. Norm Ellis, Ph.D.
External Examiner
Retired as Director
Research Division
Vancouver School Board
Vancouver, B.C.

Date Approved: 84

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

An Evaluation of a Grade Six Drug Education Program

Author:

(signature)

David H. Ericksen

(name)

17 Dec 84

(date)

ABSTRACT

A decade after the U.S. National Commission on Marijuana and Drug Abuse declared a six month moratorium on drug education in public schools, some conclusions regarding favourable outcome are beginning to appear in the literature. Most programs produce knowledge changes. Where changes in drug-related attitudes were measured, the outcome was usually positive, with some programs achieving no effect. Finally, mixed results were also seen with behaviour, the "bottom line" criterion. Contrary to popular mythology, though, increased drug use is the least common behavioural outcome.

The Grade Six program examined here, "Making Decisions: An Approach to Prevention", is one of the newer strategies. Its eight lessons include foci on advertising, peer pressure, and decision-making as well as discussion of alcohol and tobacco. Students from twenty-one experimental and five control classes completed a questionnaire before and after the lessons. Forty-nine items assessing knowledge, attitudes and behaviour were analyzed in two ways--pre/post changes in the experimental group, and experimental/control differences at post-test.

Results from both sets of analyses suggest the eight to ten lessons have achieved the immediate program objectives set by the sponsoring agency. On the first analyses, experimental students reported generally higher levels of knowledge, and more

caution towards tobacco, compared to controls. On the second (pre/post) analyses, students reported more knowledge in virtually all instructional areas, as well as more cautious attitudes in all three drug categories. No behavioural differences were apparent.

Future research, involving some alterations to the questionnaire as well as measures that go beyond the self-report method, are suggested. Similarly, program impact with Grade Five and Seven populations has yet to be assessed. Subsequent follow-up research must also assess the ultimate behavioural impact of the program over time--the most elusive and yet most desired outcome.

ACKNOWLEDGEMENTS

The author would like to thank a number of people without whom this thesis, or this degree, could not have been possible. The words in this section cannot do justice to the sentiments they represent.

First, a huge thank-you to all the dedicated people at A.D.E.S. To Frank Dingman and Norm Ellis for your great efforts over many years, to Art Steinman and May Graham who continue the task full-time, and to Colin Mangham *in absentia*, my heartfelt thanks for the trust you have extended me over the last two years.

A word of appreciation also goes to the many teachers and principals in Burnaby, Vancouver, Coquitlam and Langley who have made this project--and the program--possible. With the cooperation of your School Boards, it is your *extra* effort that has allowed this supplementary curriculum its wide exposure.

People at the Psychology Department here at SFU deserve a nod. Thanks, Bruce Alexander, for your patience and your supervisory "nudges" over the years. I've learned a lot about scientific approaches to social issues. Others in the department to whom gratitude is extended are Ron Roesch, for your capacity as committee member, and Joan Foster, for not infrequent consultations.

Finally, I am grateful to the Steel Fund at SFU, whose grant to the Senior Supervisor has made this larger- and longer-than-normal thesis research possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Since the "drug explosion" of the mid- and late Sixties, scholars, educators, and medical and legal personnel have extended great effort in mounting drug education programs. The aim has been to stop or slow the increase in quantity and variety of psychoactive substances taken by young people. Many jurisdictions, including a majority of states in the U.S.A., have required drug prevention programs in their school curricula (Blum, Garfield, Johnstone and Magistad, 1978).

Some early evaluations (e.g. Macro Systems, 1972; Swisher, Crawford, Goldstein and Yura, 1971; Stuart, 1974), however, suggested that drug education was not only a waste of time, but actually increased the use of drugs. The commercial media lost no time in dramatizing the preliminary results (Detroit Free Press, 1972; New York Times, 1972), declaring that, rather than reduce drug use, drug education stimulated students' curiosity such that they rushed out to apply their recently acquired knowledge. This "pushing or preventing" question was so common that the U.S. National Commission on Marijuana and Drug Abuse (1973) recommended a halt to all prevention programs. The

moratorium was declared not only because the outcome was undesirable, but because the research methodology of the vast majority of evaluations done at that time was so weak that the studies were, for the most part, scientifically uninterpretable.

Since that time, a large number of drug education programs (DREP's) have been offered to youngsters of all ages. To a large extent, though, the controversy regrding program outcome has not died--due in part to the still-woeful state of DREP evaluations. The purpose of this investigation is to help resolve that "pushing or preventing" controversy.

In addition to contributing to that resolution the present program evaluation addresses an immediate, applied need. In the last three years, the in-house evaluation of the DREP offered by the B.C.'s Alcohol-Drug Education Services (ADES) has been constrained by agency resources. Using only pre- and postprogram knowledge measures, together with student and teacher feedback, they were not able to develop an instrument that assessed the attitudinal and behavioural effects of the program. Neither were they able to use control groups to assess the normal developmental changes in these students.

The literature review that follows sets the programmatic and evaluative context for the needed research. The primary sources rigorously evaluate some school-based programs. In addition, four published reviews will serve to introduce methodological considerations from unpublished evaluations which were not available to me.

Three Generations of Drug Education

DREP's have three relatively distinct historical stages. Prior to the "psychedelic explosion" of the late Sixties, the few existing drug education programs relied heavily on scare tactics, typified by the celluloid classic "Reefer Madness" (Gasnier, 1936). Most programs used only one or two films, and were usually shown in the auditorium to the entire school population. Most were poorly conceived and "in direct conflict with the truth" (Baker, 1973). The strategy was to surround "drugs", which rarely included alcohol, tobacco or prescriptive medications, with revulsion, fear, guilt, shame, etc.

With the tremendous increase in drug use during the 1960's, a second generation of prevention programs emerged. They featured a straight-forward, nonevaluative presentation of the "cold, hard facts", including identification of drugs, psychological effects (both desired and undesired), toxicity, and pharmacology. But in 1972 and 1973, the "pushing or preventing" controversy, based primarily on two evaluations and two stories in major American newspapers made this strategy unpopular.

The development of the third generation of drug education/prevention programs is less historically distinct, but began after the US National Commission moratorium. The newest strategy--born out of the wish to avoid the alleged "pushing" effect of information-based preventive efforts--focuses on the

development of mastering skills, self-confidence, decision-making skills, etc. The common assumption was that the need for drugs arises from the problematic lack of those vital characteristics. That is, information alone was seen as insufficient.

Unfortunately, none of the three generations has enjoyed the scientific scrutiny necessary to render firm conclusions across populations and program variables. For example, the authors of three of four literature reviews examined below have bemoaned the lack of sophistication in both published and unpublished sources. The following review will examine the best school-based studies, with respect to problems in measurement, design, and overall research sophistication.

Primary Sources: School Based Programs

Only school-based research from which interpretable conclusions can be drawn were selected for this review, viz. research designs that used pre/post measures, and that compared treatment effects with either a no-treatment control group, or compared the effects of two or more interventions. These criteria excluded the vast majority (37 of 44) of evaluations. Three other papers will be examinined before turning to the seven methodologically adequate studies: Halpin and Whiddon (1977a) and Swisher et al. (1971) are presented for illustrative purposes; the Alcohol-Drug Education Services report (ADES, 1981) is used as the immediate context for the present research.

Table 1.1 presents the ten studies in summary form.

Swisher, Crawford, Goldstein and Yura (1971) surveyed a nonrandom sample of 993 high school and college students (age unspecified) to establish a relationship between attitudes, knowledge and use of drugs. Behavioural measures included current and total accumulated use, as well as age of first use.

Their results did not reveal any significant correlation involving drug use. They did show a significant correlation overall between knowledge and attitudes--the more students knew, the more pro-drug use was their stance. Also, marijuana users were consistently more liberal than nonusers ($p<.01$). The authors concluded, on the basis of these two correlational findings alone, that:

> an approach that relies on information alone may not be sufficient to reduce or prevent the use of drugs and, in fact, may have the opposite effect. The answer is apparently not in creating junior pharmacologists. (p. 74)

A similarly sensational discussion of meagre results follows by Halpin and Whiddon's (1977a) survey of high school seniors. They collected data on students' use of alcohol, amphetamines, barbiturates, heroin, LSD, marijuana and tobacco, but none on the presence of drug education. Use of all drugs except heroin and tobacco was significantly related to knowledge about the same drugs--from a low of .20 for alcohol to a high of .37 for marijuana. From these few behaviour-knowledge correlations, the authors concluded:

> (T)hese results support the position of the task force of the National Drug Education Association and the

TABLE 1.1 PRIMARY SOURCES: SCHOOL BASED PROGRAMS

AUTHORS	STRATEGY	SUBJECTS	PERSONNEL	DURATION	RESULTS	COMMENTS
Swisher et al. (1971)	Info Only (Correl'n design)	Private H.S., Catholic H.S., College	N.A	N.A.	1. Knowledge correlates with attitudes (i.e. the more students kne , the more pro-drug use was their stance) 2. Behaviour does not correl. with use	"No junior pharmacologists"
Malpin & Whiddon (1977a)	No program (correl'n only)	Grade 12	N.A.	N.A.	Knowledge correlates with use .20 for alcohol .37 for marijuana	"Reconsider drug education"
ADES (1981)	"Making Decisions"	Grade 6	Expert	8 hours	Knowledge gains (pre vs. post)	No control groups
Barberian et al. 1976)	1. Regular courses vs. 2. Special courses vs. 3. Assemblies vs. 4. Staff training alone vs. 5. Staff training and two of above	Grades 7 to 10	Various (categorized all existing programs)	Various	Use: no significant differences among the five program modalities	
Swisher, Werner & Herr (1972)	1. Information only vs. 2. "Process" counsellor vs. 3. "Process" counsellor and ex-addict vs. 4. "Process" counsellor and non-drug-using model	Grades 9 and 11	Carkuff counsellors	10 weeks	1. Knowledge: all modalities produce equal change. 2. Attitudes: no changes 3. Use: no changes	"Failure in implementation" (only half of counsellors met minimum performance standards)
Amendolara (1973)	Experimental vs. Control	Grade 7 (8 classes)	Regular teachers	15 weeks @ 90 min.	Attitudes: positive changes (E. males more than E. Females)	Exclusive heroin focus (unexplained)
Stuart (1974)	Info only (soft drugs) vs. Info only (hard drugs) vs. Info only (all drugs) vs. Control	Grades 7 & 9	Experts	10 weeks @ 1 hour	1. Use: Exptls higher than Controls 2. Knowledge: E. more than C. 3. Worry: E. less than C.	
Royse et al. (1982)	Info only (small groups) vs. Info only (large groups)	Junior High	Experts	3 X 50 min.	Attitudes: no changes within or between groups	
Schaps et al.(1982)	Affective vs. Control	Grades 6 + 7 (36 classes)	Expert	Ten 45 min. lessons	1. Overall: No significant effects 2. Girls: 4 of 17 indicators significantly different.	"Failure in implementation
Blum et al.(1978)	Minimal info (control; 4 hours)vs. Didactic (18 hours) vs. Process (30 hours) vs. Combined (30 hours)	Grade 6 (N=629)	Teachers and experts	2 years	Using behavioural criteria: 1. For all students, regardless of starting pattern, combined best 2. For abstainers: no difference 3. For "Low, sanctioned": Didactic best 4. For "High, sanctioned": Control best	

National Council on Marijuana and Drug Abuse. Perhaps it is time to reconsider drug education.
(Halpin and Whiddon, 1977a, p. 74)

These two articles illustrate the limited methodological sophistication upon which major drug education policies were being made in the early Seventies. Other inadequate references abound in the literature (e.g. O'Rourke and Barr, 1974; Weaver and Tennant, 1973). That their interpretations and discussions went beyond their data may have been a reflection of the demands coming from politicians, the commercial press, and the public to "do something" about the drug explosion.

Closer to home, an in-house evaluation was performed in nine of the 22 classrooms participating in the Alcohol-Drug Education Services (ADES) program in the 1980-81 school year. Much of that evaluation dealt with "process", as the program was in its early, formative stages. Written comments by the (external) program instructor, as well as teachers' and students' comments, were gathered and interpreted by ADES staff.

In a limited effort to assess program efficacy, classroom teachers were asked to administer a questionnaire to their students before and after the program. The 50-item instrument assessed knowledge gains in seven instructional areas, roughly corresponding to the eight lessons.

Results from the pre-post test, based on data from experimental students only, were expressed as knowledge gains. ADES summarized their results as follows:

1. All groups gained knowledge between pre- and posttests.
2. Six of eight groups showed significant gains.

3. The combined scores of all groups showed a significant gain ($p<.01$). (p. 18)

The author acknowledged the limitation of the research design, and pointed out that the test was important in identifying trends and pinpointing strengths and weaknesses of the "Making Decisions" program (ADES, 1981).

1. In the first of the methodologically sound evaluations to be examined here, Berberian, Thompson, Kasl, Gould and Kleber (1976) took a different approach in comparing pre- and postprogram behaviour. They categorized existing drug education efforts in thirty-three schools into five categories. The programs offered to Grades Seven through Ten students were classified as: a) drug education as part of regular school courses; b) special courses; c) assemblies; d) staff training; and e) staff training plus two of the other activities. In each of three years they examined rates of change in students' self-reported drug use for each of the five modalities, compared to that of students in schools where no program was offered.

Results from the multiple comparisons showed no pattern of significant program effects, although use of all drugs by Grade Seven students in all five categories showed a nonsignificant tendency to increase faster than that of similar (control) students not receiving instruction. The opposite was true for change rates seen in other grade levels, where small but nonsignificant differences were seen in the desired direction on the alchohol and marijuana indices. There were no differences among the five categories of program modalities.

2. Swisher, Warner and Herr (1972) also warrant a rigorous rating in methodology for use of randomization and control groups. Using a sophisticated ANOVA design, they investigated outcome among Grade Nine and Eleven students randomly assigned to four formats: a) information only in a regular health unit; b) a relationship counselling group, led by a trained graduate student; c) a "relationship" group led by the counsellor and two college role models; and d) a group led by a counsellor and two former "drug abusers". The four approaches, each stratified by low, medium and high levels of IQ, met in ten weekly sessions. Efficacy was measured by knowledge and attitudes, and a health habits scale. This study was also one of few to examine process variables: an analysis of video tapes of the sessions indicated that only half of the counsellors met minimum performance standards. Their results indicated no pre/post changes in attitudes or drug use, and all interventions produced equal changes in drug knowledge. No explanation was offered for the selection of the two grade levels.

3. Amendolara (1973) conducted a lengthy series of lessons regarding heroin with Grade Seven students in a white, suburban school. She randomly assigned the eight classes to either experimental or control conditions, and compared the pre- and postprogram attitudes and knowledge levels. No description of the 22 hours of instruction was offered.

Her results suggested that the program achieved its desired effect, although the experimental group gains were related more

closely to information items than to attitude change. In addition, more was learned by the males in the experimental group than by the females (or than the males in the control group). The author did not comment on the exclusive focus on heroin with the 13 year old population.

4. Stuart (1974) randomly assigned Grade Seven and Nine classes from two junior high schools in a university community to one control and three treatment groups. For ten weekly sessions, the experimental groups received instructions about either the lesser drugs (alcohol, marijuana, minor tranquilizers, hashish, nicotine, and coffee), the major drugs (LSD, amphetamines, barbiturates, and the narcotics), or both. Instructors were two female "experts", outside of the schools' regular teaching faculties. A three part questionnaire assessed knowledge, past and present drug use, and drug-related attitudes at pre- and posttest and at follow-up.

Stuart's behavioural results applied only to alcohol, marijuana, and LSD, as the remaining drugs were used by less than five per cent of participants, yielding insufficient variance for analysis. Content of instuction had no effect on outcome: the following results were similar for the three treatment groups. Of the three drugs reported, use levels were higher for all experimental groups than for controls. Regarding nonbehavioural measures, knowledge levels were greater for experimentals at posttest, and worry was less. Response rate from the follow-up questionnaire administration was too low

(40%) for meaningful interpretation.

5. Royse, Keller and Schwartz (1982) compared attitude changes within and between two treatment groups. Their drug education program was a series of three 50 minute presentations to either large or small groups of junior high school students. In their program, outside experts presented information regarding only "the course of the disease of chemical dependency and its effect on the family". The measures consisted of their own questionnaire (part of which was drawn drawn from Nehemkis et al, 1973), for which limited validation was provided. Differences between pre- and post-intervention attitudes towards drugs were not significant, nor were there posttest attitude differences between the two treatment groups.

6. Schaps, Moskowitz, Condon and Malvin (1982) conducted two-thirds of their DREP without discussing drugs. Typical of many of the newer "affective" strategies, six of the ten lessons examined "universal human needs" and decision-making in stressful situations before examining the drug-use context. An outside expert taught the forty-five minute lessons to Grade Six and Seven classes.

Their rigorous evaluation included random assignment of 36 classes to either experimental or control conditions. Students completed the authors' Drug and Alcohol Survey, which included 24 scales (e.g. Drug Knowledge, Peer Attitudes Toward Soft Drugs, Alcohol Involvement, Alcohol Costs, etc.) on which control and experimental groups were compared. They reported

several unsuccessful attempts at measures of decision-making skills, and did not mention assessment of the "human needs" section of their curriculum. Process evaluation included student feedback surveys, and student and teacher interviews.

Results showed weak but nonsignificant support for the program. The cohort showing the greatest impact was Grade Seven girls, where four of seventeen indicators (increased drug knowledge, decreased perceptions of favourable peer attitudes to soft drug use, and decreased personal involvement in both alcohol and marijuana use) reached the .05 level.

Interpreting their results, Schaps, Moskowitz et al. suggested that, while the evaluation itself was methodologically adequate, the overall lack of impact should be seen as a failure in implementation. Process feedback, together with the haste with which the program was conceptualized and implemented, suggested that a fair assessment of the curriculum package would only occur with a similar evaluation of a more systematic implemention of the program with a subsequent cohort.

7. The only primary source that involved both random assignment and comparison groups and that reported outcome vis-a-vis different pretest use levels, is that by Blum, Garfield, Johnston, and Magistad (1978). Their previous study had suggested:

> ...there seemed to be an active ingredient in both the didactic and process teaching styles,... the result of [which] was a differential effect by student, grade, and by class of drugs. (1978, p. 380)

This exemplary project involved random assignment of 629 Grade Six California students to two years of drug eduction in one of four modes. Students were randomly assigned to: 1) a minimum treatment "control" group involving four hours per year of instruction; 2) a didactic group, which received a total of nine hours per year of drug education in the normal lecture format; 3) a "process" group (a total of fifteen hours of instruction per year); and 4) a combined didactic and discussion format (fifteen hours per year). California law requires drug education in the classroom, which precluded a no-treatment control group for their research. A prepared curriculum was used by regular teaching faculty and by outside experts for each modality.

Questionnaire responses regarding "lifetime" use of a number of substances defined nine levels of involvement. Classification in any of the following categories was highly predictive (95%) of lower category use, i.e. if a student reported use of hallucinogens, s/he has probably used drugs that are "lower" on the "ladder" of drug use. The same step-wise progression seen below has been reported by Kandel (1975):

1. Abstains from nonmedical use of drugs.
2. Low frequency of alcohol and tobacco (less than ten occasions of a combination of substances, or any amount of only one of beer, wine, liquor, or tobacco).
3. High frequency of the sanctioned substances.
4. Low cannabis use (less than ten occasions).

5. High cannabis use (more than ten occasions).
6. Low amphetamine, barbiturate, or hallucinogen use (less than ten occasions of any combination).
7. High amphetamine, barbiturate, or hallucinogen use.
8. Low cocaine/heroin use.
9. High cocaine/heroin use.

Inhalant use was queried, but was not included in the index. All results were based on a one-way analysis of variance among the four groups; no pair-wise comparisons were reported. Behavioural queries used "lifetime" use , i.e. respsonses to the "Have you ever used .....?" formulation, as the criterion. Measures were discussed in terms of differences on usage patterns between pretest and posttest (two years).

Their results suggested that program impact depended on students' level of drug use at the start of the program. For students starting in Pattern One (abstainers), no differences in outcome were seen among the four groups. For students in Pattern Two (low use of alcohol or tobacco), though, those exposed to the didactic mode did best. For students in Pattern Three, viz. those already regularly indulging in alcohol and tobacco, a minimum of drug education was best. Students starting the program in Pattern Four or above (any use of the illicit substances) were too few to warrant analysis. Blum et al. also derived a summary measure for all groups, using a weighted combination of starting patterns. Using this criterion, the combined discussion-didactic mode was best for all students,

regardless of starting pattern.

Blum et al. also performed a regression analysis to find factors that predicted posttest scores. They found that the best predictors were beginning drug use, the use of inhalants, and gender (being female predicted lower scores). All variables examined in the Blum et al. study accounted for 42% of the variance in drug score changes. Only 5% of the overall variance was explained by two year educational experience.

An analysis of the teacher effects showed that significant results were limited to "outside experts", in this case members of the research staff. Greatest variability of the teacher effect was found in the discussion format, where no standardized curriculum was present to minimize individual differences.

The authors attempted two long term follow-up surveys. The first, tracing their first Grade Six cohort after six years, resulted in an insufficient response rate (21%). The second, attempting to engage a four year follow-up of their earlier project, also found sample attenuation too great.

The articles reviewed here present a pattern vis-a-vis outcome similar to that seen in the published reviews. Most outstanding is the large number of articles and citations collected in order that even a small number of methodologically adequate reports appear.

The school-based programs whose evaluations have been presented by no means show a consistent pattern. Of the various outcome criteria reported, two of the three studies which

documented knowledge changes reported positive outcome. Attitudes fared more poorly: only two of four studies reported positive outcome, with the balance reporting no effect. Regarding the behavioural bottom line, three programs failed to register an effect, while a fourth reported increased use levels for experimentals. A fifth reported mixed results, based on different patterns of use at the program's start.

Review Articles

The four review articles all concluded a lack of systematic rigour in the area. Reviewers found insufficient sample sizes, a lack of even nonequivalent comparison groups, and/or a failure to use behavioural measures. The latter is particularly disconcerting as most of the studies that did report drug use found very poor correlations with attitudes or knowledge (cf. Braucht, Brakarsh, Follingstad, and Berry, 1973; Cialdini, Petty, and Caciappo, 1981; Eagly and Himmelfarb, 1978; Ebel, Katz and Rosen, 1975; Halpin and Whiddon, 1977b; Thornburg, 1980). The entire area regarding the relationship between drug knowledge, attitudes and behaviour remains largely unaddressed in the evaluation literature.

Braucht, Follingstad, Brakarsh and Barry (1973) published one of the earliest reviews in the area. Their primary conclusion, as mentioned previously, regarded methodology:

> A major portion of the literature is devoted to ... considerable controversy, almost none of which is based on systematic empirical evidence. The following review of opinions [is thus] based primarily on anecdotal or

uncontrolled case studies...(p. 1283)

Their manuscript, prepared in 1972, was very early in the "second generation" of drug education programs. Only three of their twenty-seven references showed publication dates prior to 1968.

Braucht, Follingstad et al. (1973) reported that these formative years passed without any clear-cut philosophy concerning aims, goals and approaches, mostly because of varying attitudes vis-a-vis teenagers and drugs. For example, program objectives sought to: 1) prevent all use of all drugs, including alcohol; 2) present only information regarding identification, pharmacology, toxicities, etc. of illicit substances, thus explicitly avoiding the use/abuse distinction; or 3) discuss all factors involved in any use of licit or illicit substances (peer pressure, psychological factors, social conditions, legalities, etc.)

The bulk of the studies examined by Braucht, Follingstad, et al. (1973) were methodologically inconclusive. Four citations used anecdotal reports or sample sizes of fewer than ten. Two of the remaining studies had program durations of less than five hours. The only study discussed in any detail measured only post-intervention levels of knowledge and attitudes (but not behaviour) among volunteer Catholic students. The only program component that warranted its mention was a follow-up procedure after six weeks.

Randall and Wong (1976) also found that the vast majority of prevention efforts were poorly evaluated. Their review was

more ambitious than the one above: they collected 200 published evaluations of DREP's, of which only twenty-three included "any systematic evaluation". Fifteen citations included both comparison groups and pre- and post-intervention measures.

Of those fifteen, the six which were directed at elementary and/or high school populations were unpublished. Their duration ranged, with one exception, from four months to three years. A variety of programs compared: several "values" strategies, all of which used the lecture format; several kinds of counselling techniques (in a classroom discussion format); and moral development with lecture format.

Methodologically, only one of the fifteen citations included any measure (viz., number of drug-related hospital admissions) which used other than a self-report method. Given that the principal conclusion of their review regarded the poor methodology used in most evaluations, it seems curious that Randall and Wong did not comment on this over-reliance on paper-and-pencil measures. They do, however, comment that the vast majority of articles seen by them paid little attention to which particular variables contributed to outcome effectiveness (e.g. age, teacher/"expert", strategy, duration, etc.).

Outcome, shown in Table 1.2, was positive in all six classroom-based programs. In some cases, the authors report that not <u>enough</u> gain in knowledge or attitudes was seen, but in none of these six was there a lack of effect or a negative outcome. Where Randall and Wong reported knowledge-based measures,

TABLE 1.2. School-based DREPs from Randall and Wong (1976)

Authors	Strategy	Subjects	Duration	Results
Virgilio (1971)	Lecture vs. lecture/discussion	Senior high school students	3 weeks	Attitude and knowledge: both significant, but marginal, improvements
Warner, Swisher & Horan (1973a)	Values and decision-making vs. controls	Gr. 5 to 10	?	Use: decrease Values: positive
Warner, Swisher & Horan (1973b)	Behavioural counselling vs. cognitive counselling vs. placebo counselling vs. control	Grade 9	?	Attitudes: pos. changes for behavioural group
Warner, Swisher & Horan (1973c)	Facts vs. Values vs. Combined vs. Controls	?	3 years	Attitudes: positive change for Values and Combined modes Use: descreases for Values and Combined modes
Carney (1971)	Several Values and Decision-making modes	Gr. 4 to 12	3 years	Values: positive (e.g., higher perceived risk, lower gain)
Briskin (1974)	Moral development vs. lecture	Grade 6	7 months	Knowledge: both gain equally Communication skills: no change Moral development: no change

knowledge gains were seen. A similar pattern is seen for the four studies that documented attitude changes, as well as the two that examined drug-using behaviour. The Randall and Wong review, then, makes a reasonably strong case for positive outcome (at least in knowledge and attitudes) in medium-to-long classroom-based programs.

Drawing heavily on unpublished and abstracted evaluations, Hanson (1982) reviewed the behavioural outcome of school-based programs (summarized across program strategies, durations methodologies, personnel, etc.). Of twenty-four programs that are of interest here, six evaluations suggested reduced drug use (or, a rate of increase slower than that of controls). Ten of the studies cited by Hanson reported no behavioural change, and three reported negative results. Five others reported mixed (positve and negative) outcome.

Schaps, Churgin, Palley, Takata and Cohen (1980) used the benefit of time and accumulated wisdom to summarize thirty-five of the most rigorous evaluations in the literature by types of prevention strategy, institutional setting, age range of target population, rigour of research design, type of impact measure--and categorized outcome as either positive, negative, mixed or no impact.

In terms of prevention strategies, Schaps, Churgin et al. (1980) noted that the trend in the last five or six years is towards evaluating multi-dimensional programs, i.e. those combining informational, affective (values clarification,

decision-making skills, communications skills, etc.) and/or peer-involved approaches. The authors noted the total absence of family-oriented strategies in that time period. Table 1.3 shows that only 40% of the 35 studies cited used the didactic procedure alone--a major change since Braucht, Follingstad et al. (1973). Target populations were predictably school children --two thirds of the evaluations were directed at captive minors. About a quarter of all studies were aimed at eight to eleven year old children, most of which were published since 1973. More than a third assessed effects on twelve and thirteen year olds. Sixty per cent examined high school student (14 and 15 years) populations [1]. Most used only paper-and-pencil methods. Only one of the thirty-five used two independent measures (student self-report and teacher estimations) in measuring drug use. Three studies used law enforcement statistics or hospital records.

Outcome results overall appeared better for the programs that involved an affective component than for the older style Information Only programs. Table 1.3 shows reductions in drug use in about half the Affective Only studies and in about one third of the programs involving an affective component. Contrarily, no Information Only programs reported positive effects on behaviour. Four of seven programs reported no effect, with the remainder reporting negative or mixed results.

---

[1]Total exceeds 100% because some studies compared multiple age groups.

TABLE 1.3. Program Outcome by Strategy

Strategy	No. of Studies	Outcomes: Behavioural Measure	Outcomes: four non-behavioural measures
Information only	13	4 of 7 no change 1 of 7 negative 2 of 7 mixed	6 some positive 6 some negative
Info/Peer	3	1 negative 2 no effect	
Affective Only	5	half* positive	17 of 22 some pos. 7 of 22 some neg.
Info/Affective	11	one-third* positive	
Info/Affective/Peer	6	one-third positive	
Other	8		

*Inconsistent reporting throughout body of paper.

Combining four nonbehavioural measures (decision-making skills, school attendance, knowledge, and family functioning), the Information Only strategies demonstrated some positive effects in six of thirteen studies, but negative effects in the same number. The newer generation, involving an affective component, got a positive rating in seventeen of twenty-four studies and showed negative effects in seven.

Schaps, Churgin et al. warned that the novelty effect involved in the newer strategies may account for their greater

efficacy. Similarly, level of rigour involved in the new strategies as evidenced by the presence of a control group, may not be as great as with older approaches. The authors cross-tabulated the four levels of outcome (positive, negative, mixed, none) with level of rigour and found that as the rigour of research design increased, the likelihood of negative effects increased and the likelihood of positive effects remained fairly constant.

Methodological Issues

Methodological inadequacy is not unique to DREP's. Gilbert, Light, and Mosteller (1975) described evaluations of federally funded manpower training programs in the United States, on which $6.3 billion were spent on 6.1 million people between 1963 and 1971. Citing a U.S. National Academy of Sciences (1974) report, they wrote:

> Manpower training programs have been in existence a little over a decade, yet .... little is known about the educational or economic effects of the programs. This is troublesome, in light of the fact that about $180 million have been spent in an attempt to evaluate them. (in Gilbert et al, 1975, p. 161).

School-based smoking and health programs, which predated DREP's by six or eight years, show the same pattern. Since the impetus for smoking prevention programs and their evaluations came from the first U.S. Surgeon General's report (USDHEW, 1964), the early haste in programs and their evaluation can be historically understood. Similar to the DREP literature, though, adequate research design in smoking and health programs is still

the exception rather than the rule (Evans and Raines, 1982; O'Rourke, 1980; Thompson, 1978).

In 1973, the six month moratorium on DREP's in the U.S. was as much a reflection of the lack of methodological sophistication of DREP evaluations as of the program outcomes per se. While limited improvements in knowledge were attributed to DREP's, for the most part evaluations provided "little scientific evidence from which one could confidently draw conclusions" (Blum, 1976, p. 51).

Design

A recent review by Schaps, DiBartolo, Moskowitz, Palley and Churgin (1980) found that there was still much room for improvement. In terms of research design, they reported that 58% of the 127 published and unpublished evaluations reviewed used pre- and post-intervention measures, while 25% used only a posttest. Only 12% included a follow-up component.

Their findings regarding the quality of comparison groups was instructive. Group "equivalence" was found to be adequate (defined as of sufficient comparability to yield clearcut and definitive findings) in only 40% of cases--31% used random assignment and 9% used strongly equivalent nonrandom groups. The remaining 60% used either weakly equivalent (28%) or unacceptable (15%) comparison groups or none at all (17%).

Reviewing the strength of the research designs as a whole, Schaps, DiBartolo et al. rated 31% as unacceptable, 28%

borderline, 21% acceptable, and only 20% as strong. No improvement in the quality of design was seen over the decade covered by their review. Program outcome (positive, neutral, or negative) was not related to overall research design quality, nor with quality of comparison groups.

Measures

A quarter century ago, Campell and Fiske (1959) published a classic and now time-worn prescription regarding the use of different methods and different traits in measurement. Their "multimethod-multitrait" prescription held that virtually air-tight validity could be demonstrated by the use of independent measures of more than one trait. Unfortunately, their advice has been largely ignored in drug education research.

Ninety-two per cent of the evaluations reported by Schaps, DiBartolo et al. used only a single method, viz. self-report. Furthermore, the vast majority used only a questionnaire. This ubiquitous instrument, used alone in 88% of all studies, was accompanied by interviews in 12% and archival information in 8% of cases.

Attitudes were the most commonly measured trait (76%). Schaps et al. reported that 54% of evaluations asked about the use of drugs, and 43% inquired about knowledge.

An examination of testing effects in the area has barely begun. Using a Solomon four-block design, Casswell (1982) found

that control groups who completed both a pre- and a posttest reported more alcohol use, drunkenness, and marijuana use, compared to control groups who completed only the posttest. She also found an interaction between pretesting and the experience of the DREP itself. To the degree that pretesting sensitizes subjects and/or acts to allay their anxiety, DREP researchers must seriously consider elaborating both their design (beyond the pre/post control groups considered minimal here) and their measures.

Little can be said regarding the validity of the measures used in the area, because so many authors have failed to report validation procedures. Likewise, there has been little attempt to develop a standardized questionnaire or pool of items (Hochhauser, 1979). Finally, issues regarding the validity of questionnaires (either anonymous or identified) compared to interviews have not been discussed in regard to classroom-based DREP's.

Single, Kandel and Johnson (1975) assessed under-reporting of drug use by students over multiple questionnaire administrations. They found that inconsistent responses--where a student reported less use of a particular substance on a later administration--increased over time. Those inconsistencies usually related to only one substance with which the student had had only passing experience. Under-reporting, they concluded, was a function of poor recall rather than deliberate concealment.

A number of authors have examined the opposite problem--over-reporting. Overall, less than 1% of the population in the Single et al. study indicated use of "adrenochromes (wagon wheels)". Similar proportions were reported by Petzel, Johnson and McKillip (1973), where 4% of their high school sample endorsed use of "bindro (stars, hexahydrol)". Likewise, Halpin and Whiddon (1973a) concluded that over-reporting is not a problem for the questionnaire method, as only 1% of their sample acknowledged use of "CHD".

Finally, a remark to those who would prefer a more "objective" measure, rather than self-report. While many authors have dismissed the use of blood and/or urine samples on ethical grounds (e.g. too invasive), Single et al. (1975) cited a report that examined the technical virtues of such methods. Examining the results reported by over 1500 commercial laboratories, Berkowitz (1974) found that only 64% "correctly identified most of the drugs include in the sample" (in Single et al, 1975, p. 441).

Almost all of the DREP evaluations have been based on measures taken immediately following the intervention. Few have assessed program impact over time, e.g. only one of the 200 reports cited by Randall and Wong (1976) included a follow-up component. Similarly, only three of the 45 articles collected by this author reported follow-up data. Two of the three (Stuart, 1974; Blum et al., 1978) found response rates too low to allow meaningful interpretation; the third found that positive

attitude changes and reduced tobacco use at posttest were not evident a year later.

The demand for more research can be expected to come from academics and social scientists. All the requirements for increased rigour--for better design, for better measures that incorporate multiple traits and methods, for better access to recipients of program services, etc.--come not simply from striving for unending methodological perfection but from the "bottom line" criterion. Gilbert et al. (1975) suggest that the degree of research rigour which DREP evaluations must meet is the answer to a difficult question: "How much should society be prepared to pay for an evaluation of a program that leads to firm and reliable conclusions about how well the program is working?" The corollary, however, may be more easily answered: "How much should be paid for an evaluation that establishes that a program is <u>not</u> working?"

II. METHODS

The program being evaluated is one sponsored by the Alcohol-Drug Education Services, an independent community agency that offers a number of educational services in the Lower Mainland of British Columbia. Their program, run in Vancouver, Burnaby, and Langley schools since 1979, has used a decision-making strategy where lessons about advertising, concern for others, peer pressure, decision-making, and communication skills receive more attention than the drugs themselves. In the 1983-84 school year, some 115 classes received their curriculum. The program continues to expand.

From the beginning, ADES has run a small scale evaluation based on a short questionnaire administered before and after the program, to students receiving the eight lessons. The early questionnaire, measuring only changes in knowledge that pertained to curriculum content, was accompanied by a "process evaluation" that asked for student and teacher feedback. In May, 1982, ADES kindly consented to the present evaluation, which would utilize more students, comparison groups, and a more detailed questionnaire based on behavioural, attitude, and knowledge measures.

Until the 1982-83 school year, ADES staff have taught the lessons in all the classrooms. Beginning in January, 1983, the focus for the "decision-making" program has been to orient and support the classroom teachers in presenting the materials. With this strategic shift, forty-five classes in Vancouver, Burnaby, Richmond, and Coquitlam received the program. ADES staff continue to provide a parent night for each school prior to the commencement of the unit.

Subjects

Participation in the evaluation, as in the program itself, was voluntary. Some forty-five Grade Five, Six and Seven classrooms in Burnaby, Vancouver, Coquitlam, and Richmond were taught the lessons in 1982-83, mostly between March and June. Of that number, only one declined the evaluation component.

The first twenty-one classes to notify ADES of their imminent commencement served as the experimental group. Pre- and postprogram data were collected from sixteen of the twenty-one classrooms. One teacher did not finish due to prolonged illness, another finished too late in the school year to allow posttesting, and three others did not begin the program (reasons unknown).

Teachers' timetabling did not allow control groups to be constituted as originally proposed. The "waiting list control" variation of the pretest-posttest control group design was not possible: teachers did not provide starting dates sufficiently

in advance to allow the later-starting classrooms to serve as controls for the earlier experimental groups. Therefore, teachers who had made a commitment to use it the following year were asked to make their classes available as control groups. Control data from five Grade Six classrooms were gathered in May and June, 1983.

Parental consent procedures produced minimal sample attenuation. Less than thirty experimental students (7%) returned negative consent forms. Parents in the control groups were even more cooperative--only 5% declined. The few students who failed to return consent forms were instructed by their teacher to withdraw from the evaluation.

Geographically, the experimental and control groups were dispersed throughout the Lower Mainland. Three classrooms from two schools in Burnaby, eleven classes (four schools) in Vancouver, and two classes from Coquitlam (two schools) comprised the experimental group. Two classes in one Langley school, two in one Burnaby school, and one Vancouver class served as the control group.

The proportion of girls in the complete data set was 60% for the experimental group and 53% for the control groups. Of the cases used in the analysis (see "Case Selection" below), 53% were girls in both the experimental and control groups. The experimental students were an average three months older than controls (11.94 vs. 11.69 years; $\underline{t}(102)=4.15$, $p < .05$).

Instrumentation

The questionnaire used in this evaluation (Appendix II) has tried to maximize the use of items from previous research. The behavioural items are a composite of those used in Einstein and Allen (1972) and Blum et al. (1978). The knowledge portion came from last year's in-house evaluation of the "Making Decisions" program (ADES, 1981), and from various instruments reprinted in Cornacchia, Bentel and Smith (1973) and Nehemkis, Madari and Lettieri (1976). Items were selected for their relevant content where the sentence structure and grammar was at a Grade Six level.

The nine attitude questions were selected from the Drug Attitude Scale (Goodstadt, Cook, Magistad and Gruson, 1978), about which validity and reliability statistics have been reported.

Procedure

About two weeks before the program began, classroom teachers briefed their students about the upcoming program, and asked them to return parental consent forms. For the experimental classes, the consent forms included announcement of a parent night, which was normally conducted by ADES staff in the week prior to the program's commencement. The attendance averaged ten to fifteen parents.

On the day the preprogram questionnaire was administered, within a few days of commencement of the DREP, teachers directed

nonparticipating students to another classroom or the school library. In most cases, teachers remained in the classroom for the duration of the procedure (thirty minutes).

The researcher introduced the evaluation by paraphrasing the "student information sheet" (cover of questionnaire). The ususal emphasis was placed on confidentiality, the voluntary nature of their participation, and the importance of honesty. Instructions regarding the "secret code" were given one step at a time: e.g. students filled in the first digit before receiving instructions for the next digit. Questions of clarification were answered; no explanation of terms was offered for knowledge items.

Most students took twenty minutes to complete the questionnaire, with rare exceptions taking thirty minutes. Students occupied themselves with reading or other quiet work for the remainder of the testing period.

The only procedural difference for the control classrooms arising from the research design itself was an explanation of the need for comparison groups. Three or four control students, and no experimental students, expressed doubt about that rationale, e.g. "Why are you <u>really</u> doing this?"

Three of the five control classrooms were particularly rambunctious during the questionnaire administration(s). On one occasion, a classroom had just returned from a vigorous and "heated" game of soccer. Two other classrooms in one school were combined for one administration--resulting in a boisterous

reunion of peer groups which had been deliberately separated at the beginning of the school year. With all control groups, the classroom ambience was more restless than that of the experimental groups.

Research Design

Two questionnaire administrations to sixteen experimental and five control classrooms comprised a pretest-posttest control group design. Classrooms were not randomly assigned to conditions: rather, pragmatic considerations dictated that the first group of teachers to start the program would serve as the experimental group. Similarly, the five classrooms which served as control groups were selected because the teachers had expressed interest in offering the program in the 1983-84 school year.

The pretest-posttest control group design used here is one of the most popular research designs in clinical and educational research (Campbell and Stanley, 1963; Cook and Campbell, 1978; Kazdin, 1980). With random assignment, it provides strict control over threats to validity in assessing the effects of different interventions.

Many research problems, including the present one, do not allow the randomization of individuals to groups, or groups to treatments. Generally interpretable results are possible where intact groups preclude assignment of individuals to groups (Campbell and Stanley, 1963; Evans and Anastasio, 1968), if care

is taken to ensure systematic differences are not present at pretest (cf. "Results: Preprogram Comparisons", below).

Statistical Analysis

Preprogram equivalency was tested by running a series of group comparisons of pretest scores. A similar, and statistically unrelated, analysis of posttest scores provided an overall account of program effects. Sample size was relatively small here. The usable control data numbered 52 cases, and a random sample of the experimental data were selected to provide equal cell sizes (see pp. 40-44 for details on case selection).

To take advantage of the larger data set, simple pre- vs. post-program comparisons were performed using only the experimental cases. The approximately 220 questionnaires in each administration provided for a substantial increase in degrees of freedom. While the pre/post-only design has weaknesses when used as the only basis for comparison, it complements the experimental vs. control posttest comparisons fairly well.

For both sets of group comparisons, the Yates-corrected Chi square tested for differences in proportions in experimental and control groups (Dixon, 1981). Where the mimimum estimated expected value for a cell in the Chi square test was less than 5, the test was disregarded (Dixon, 1981, p. 157). The basic T-test was used in comparing differences in group means unless Levene's test for equal variance suggested heterogeneity, in which case the Welch-adjusted T-test was used (Myers, 1981).

Finally, use of a more precise statistical model was also explored. Analysis of covariance, using individual students' pretest scores as the covariate, can be used to increase the precision of the overall F test(s). This procedure assumes that pre- and posttest scores are not independent: hence, statistically "adjusting" for that systematic component of the posttest score that was present prior to the intervention reduces the error term (the denominator in the F ratio).

Some authors have suggested the use of repeated measures ANOVA (with pretest scores as the within factor) or ANOVA of gain scores to examine outcome. Using the present design, though, the two approaches would be too conservative, or misleading, or both. Huck and McLean (1975) have advised that the gain scores approach rests on the same set of assumptions as does a regular ANOVA--the only difference is that the regression (Beta) weight is arbitrarily preset to 1.0. Further, the gain scores analysis does not allow identification of regions of significance or nonlinear regresssions. Finally, estimated treatment effects using a repeated measures ANOVA will be spread out over the pretest scores, causing an F ratio that is conservative by half. The authors wrote:

> ... the interaction F in repeated measures ANOVA is the one that is really dealing with the main effect of treatments..... The interaction F and the F from the gain scores analysis will always turn out to be identical, regardless of whether the pretest means are equal to one another. (Huck and McLean, 1975, p. 513).

Some arguments have been made against the use of ANCOVA in certain applications of the pre/post control group design. For

example, Gourlay (1953) strongly advised against it when groups aren't randomly assigned, even if the matching of predata has produced no significant differences. If there are significances at posttest, he suggested:

> ...there is no guarantee that there may not be other differences between the groups operating to produce final differences which will be unattributable to the program. (1953, p. 30)

Similarly, where intact groups are used but treatments are assigned to groups at random, Elashoff (1969) wrote:

> We can never be sure that the covariance has removed all the bias--some bias may still be present from a disturbing variable which was overlooked. (p. 386)

Such critiques, however valid, are not properly addressed to the ANCOVA statistical analysis. They are questions of design--of the statistical equivalence of nonrandomly-constituted groups. The above criticisms can be made equally to analyses of variance of posttest scores alone. Comments regarding ANCOVA are fair only if a researcher does not offer evidence regarding the <u>additional</u> assumptions regarding homogeneity and linearity of regression.

III. RESULTS

Results from two sets of analyses are presented. First, the stronger design--post-program comparisons of experimental vs. control data--is presented for the groups as a whole and separately for girls and boys. The second set of analyses, based on pre- vs. post-program data from the experimental group only. Brief comments on analyses of covariance are also provided, as are procedures assessing validity of the questionnaire.

Subscale and Index Construction

Many of the forty-eight questionnaire items have been summed to provide indices of substance use, attitudes toward and knowledge of alcohol, tobacco, unspecified drugs, and curriculum-related knowledge.

The drug use index, ALLUSE, was derived by Blum et al (1978). A detailed description of the categories has been provided above (pp. 13-14). Briefly, the levels of use are: 1) abstention; 2) low levels of use of the sanctioned substances (alcohol and tobacco; 3) high use of the sanctioned substances; 4) low use of the soft drugs; 4) high use of the soft drugs; 6) low and 7) high use of pills or hallucinogens; and 8) low or 9) high use of cocaine or heroin.

Students were scored as "abstainers" (Pattern One) if they did not report use of alcohol or tobacco, or specify use of any drugs in the "Other--please specify" item. If only one of the

cigarette or alcohol items (23 or 24 - 26, respectively) were endorsed at any level, a student was placed in Pattern Two (low use of the sanctioned substances). If either alcohol or tobacco had been used more than ten times ("f", "g", or "h" on items 23 - 26) in the presence of any use of the other, a "high--sanctioned" rating (Pattern Three) was assigned. A student who specified use of marijuana or hashish on less than ten occasions received a Pattern Four rating. Similarly, more than ten occasions rated a Pattern Five score. On only two or three occasions were other psychoactive substances specified in the "other" item. Endorsation of item 28 without specification was not included in the ALLUSE index.

The attitude subscales and overall index are taken from Goodstadt et al. (1978). Five part Likert-type scores from each of three items were summed to provide the corresponding attitude subscale: unspecified drug attitudes (items 46 - 48); tobacco attitudes (items 43 - 45); and alcohol attitudes (40 - 42). The raw scores from all nine attitude items were summed to provide the overall attitude index.

The three drug knowledge subscales were defined by the respective number of correct answers: unspecified drug knowledge included items 8, 13 and 49; items 7, 9, and 10 constituted the alcohol knowledge subscale; and correct answers to 5, 6, 11, 14 and 15 defined cigarette knowledge. Similarly, the following curriculum-related knowledge subscales came from items 29 - 33 (Parts of the person), 34 - 39 (Wants vs. Needs), 16 and 17

(Advertising), 18 and 19 (Decision Making), and 19 and 20 (Peer Pressure). For both knowledge areas, drug knowledge and nondrug knowledge, an overall index reflects the grand sums. The final index, All Knowledge, is the total number of correct answers from all drug and nondrug items.

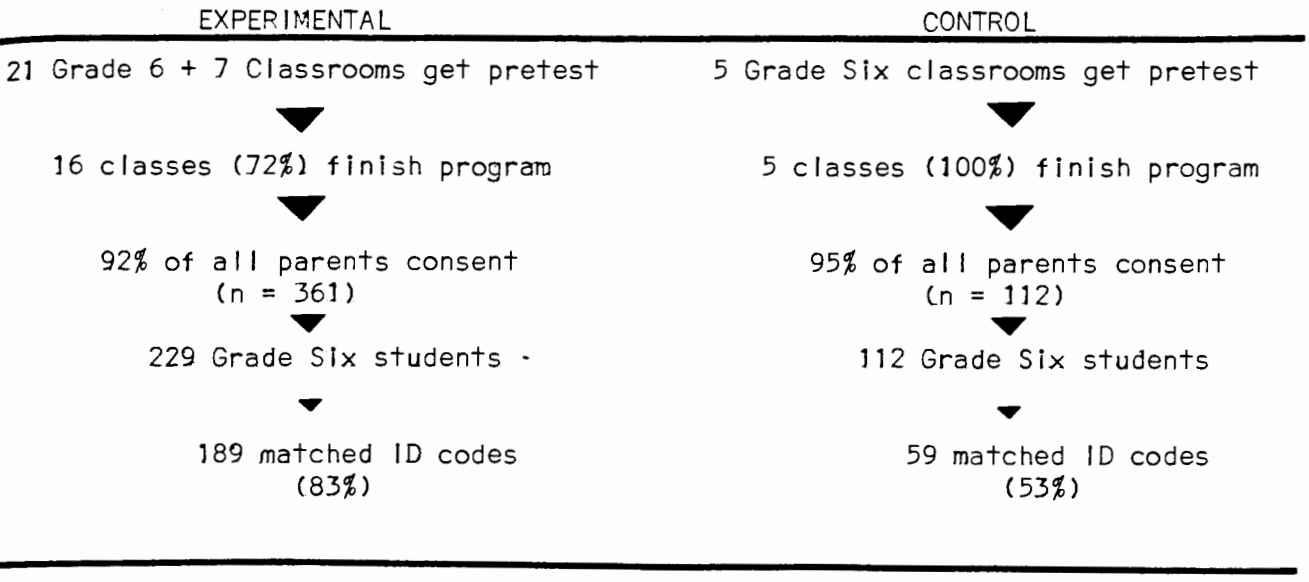
Experimental vs. Control Postprogram Comparisons

Case Selection

Not all of the subjects could be used for the pre- and postprogram comparisons. Figure 3.1 shows that 361 students in Grade Six and Seven were to serve as the experimental group, compared to 112 control students. Unfortunately, due to late changes in the research design, no Grade Seven control classes were available. Experimental data from 132 Grade Seven students were therefore excluded. As such, the present evaluation is not of the ADES program as a whole, but only of its impact on Grade Six students.

One further, and serious, sample attenuation remains. Preliminary descriptive statistics showed a substantial disparity in proportions of experimental and control students who were present, able and willing to transcribe the same "secret identity code" on each of their two questionnaires. The proportions of these "paired" questionnaires were 83% for experimental, and 53% for control, students.

FIGURE 3.1 Sampling Procedure



The failure to provide "paired" questionnaires could occur for a number of reasons. First, about five per cent of students were not present on any day to participate. Absences due to illness or doctors' or dentists' appointments would account for approximately a 10% loss in the "pairing" rate for the two questionnaires, leaving 90% of students present for both administrations. The observed experimental rate, then, suggests that most of those students present were able and willing to participate.

The relatively low "pairing" rate for control groups (53%) remains without satisfactory explanation, but two observations may shed some light on the problem. First, there were slight differences in protocol in the in-class administration of the questionnaires: the rationale for control groups in research

design was not readily understood by the twelve-year olds. The plausibility of the procedure may not have been as high as was hoped--control students had not had the legitimizing influence of a parent night, nor been part of the program itself. Their only experience was of a visitor who introduced himself as a researcher, who claimed the information was for scientific purposes. In all control classrooms, the overall restlessness may be a sign that a proportion of control students were not willing to transcribe the same ID codes.

Secondly, some control students may not have been able or willing to partipate because of extraneous factors, unique to these control classes. As mentioned previously, in three control classes, students were quite rambunctious and/or aroused due to a soccer game immediately preceding the administration in one class, and because of a boisterous reunion of peer groups in two others.

To determine the effects of the planned and unforeseen differences in protocol, the "paired" vs. "non-paired" control data were examined for group differences. Table 3.1 shows that the two data sets could not be considered equivalent. Significantly more of the "non-paired" questionnaires showed use of cigarettes as defined by item 23 (74% vs. 41%; Chi Square=10.41, $p < .01$). Similarly, significantly more nonmatched students reported use of liquor (62% vs. 39%; Chi Square=4.72, $p < .05$). As well, each of the five remaining specified behavioural items (regarding either tobacco or alcohol) was endorsed by a

substantially, but not significantly, higher proportion of nonmatched students. Significant differences in mean attitudes toward drugs (unspecified) were evident: nonmatched students were more conservative (7.81 vs. 8.69; $t(108) = 5.65$, $p < .05$). Finally, the nonmatched controls did significantly better on the "Wants vs. Needs" knowledge items (1.56 vs. .98; $t(78) = 5.71$, $p < .05$). Given these four significant differences, and the trend toward more nonmatched students reporting drug use, the two groups of control data could not be seen as equivalent.

Given the nonequivalency of the matched and nonmatched data from the control classrooms, the matched control data were selected for comparison with the experimental group because the matched rate was so high in the experimental group (83%). Similarly, of the two control data subsets, the more valid information should come from students who took enough care and who were trusting enough to transcribe the same ID code. If that is the case, the matched control data should be equivalent to the matched experimental data at pretest.

Assessing Preprogram Equivalency

Analysis of the preprogram data (Table 3.2) suggests roughly equivalent patterns of use, attitudes and knowledge between the experimental and control groups prior to intervention. Pre-program use of tobacco was queried in two ways: about one third of each group reported either any use at all (item 23), or that they had tried "more than a puff or two"

TABLE 3.1 PREPROGRAM DATA: MATCHED VS. NON-MATCHED CONTROLS

TABLE 3.1a CATEGORICAL DATA

No.	Variable	Matched Controls	Non-Matched Controls	D.F.	χ^2 *	Prob.
1	SEX: % Female	53	53	1	.00	
21	CIGARETTE USE: % Yes (more than a puff or two)	31	46	1	2.03	
22	ALCOHOL USE: % Yes (e.g., a bottle of beer, without consent)	13	30	1	3.68	
23	CIGARETTE USE: % YES (any, with or without consent)	41	74	1	10.41	p < .01
24	BEER USE: % Yes (any, with or without consent)	74	84	1	1.02	
25	WINE USE: % Yes (any, with or without consent)	73	86	1	1.86	
26	LIQUOR USE: % Yes (any, with or without consent)	39	62	1	4.72	p < .05
27	BOGUS USE: % Yes	0	18	N.A.[1]	N.A.[1]	
28	OTHER USE: % Yes (if specified)	4	4	N.A.[1]	N.A.[1]	

*Yates corrected χ^2 [1]Minimum estimated expected value less than 5.

TABLE 3.1b CONTINUOUS DATA

Variable	Matched Controls	Non-Matched Controls	D.F.	t ratio	Prob.
AGE	11.69	11.75	110	.23	
INDEX OF USE ("All use")	2.17	2.27	97	.94	
ATTITUDES					
Drugs (unspecified)	8.69	7.81	108	5.65	p < .05
Alcohol	9.32	8.92	107	.87	
Cigarettes	6.20	6.35	101	.11	
All Attitudes	24.20	23.10	101	1.74	
DRUG KNOWLEDGE					
Drugs (unspecified)	1.28	1.43	94	1.28	
Alcohol	1.24	1.12	108	.76	
Cigarettes	2.73	2.73	108	.00	
All Drug Knowledge	5.32	5.27	91	.02	
NON-DRUG KNOWLEDGE					
Parts of the Person	3.59	3.53	79	.89	
Wants vs. Needs	.98	1.56	78	5.71	p < .05
Advertising	.80	.69	106	.61	
Decision-making	1.04	.88	107	1.46	
Peer pressure	1.02	1.10	109	.51	
All Non-drug knowledge	7.67	8.06	66	.59	
ALL KNOWLEDGE	13.28	13.24	61	.00	

(item 21). An equal number of each group (11% and 13%) admitted consuming an alcoholic beverage without parental consent. Differences in any use of wine approached but did not reach significance--54% vs. 73% (Chi Square=3.13, $p < .07$). were reported by the experimental and control groups. Use of liquor by all students was lower--69% and 62% (respectively) had not tried spirits.

The classification pattern of drug use was intended to define abstinence, low or high use of the sanctioned substances, low or high use of the "soft" drugs, etc. The vast majority of these groups of these eleven and twelve year olds, however, fell in the "low--sanctioned" category. The two groups, with means of 2.06 and 2.05, did not show homogeneity of variance (latter not shown). An examination of the histograms for the two groups (not shown) reveals that only one of the 45 students reported trying anything more than alcohol or tobacco on ten occasions. Likewise, only six of 41 control students were "nondabblers".

Table 3.2 also shows the three attitude subscales and the overall drug attitude index: not only is there a lack of significant differences between groups, but pre-intervention scores in each case are virtually identical. Notable, though, are differences between subscales: where a summed score of 3 implies very cautious views, a score of 9 suggests neutrality, and 15 indicates little caution towards drugs, alcohol was regarded neutrally (group means of 9.26 and 9.32) while tobacco attitudes were decidedly conservative (6.78 and 6.20).

TABLE 3.2 PREPROGRAM COMPARISONS: EXPERIMENTALS VS. CONTROLS

TABLE 3.2a CATEGORICAL DATA

No.	Variable	Exptl (n = 52)	Control (n = 52)	D.F.	χ^2 *	Prob.
1	SEX: % Female	60	53	1	.23	
21	CIGARETTE USE: % Yes (more than a puff or two)	27	31	1	.05	
22	ALCOHOL USE: % Yes (e.g., a bottle of beer, without consent)	11	13	1	.00	
23	CIGARETTE USE: % YES (any, with or without consent)	34	41	1	.24	
24	BEER USE: % Yes (any, with or without consent)	58	75	1	2.39	
25	WINE USE: % Yes (any, with or without consent)	54	73	1	3.13	
26	LIQUOR USE: % Yes (any, with or without consent)	31	38	1	.31	
27	BOGUS USE: % Yes	0	0	-	N.A.[1]	
28	OTHER USE: % Yes (if specified)	2	4	-	N.A.[1]	

*Yates corrected χ^2 [1]Minimum estimated expected value less than 5.

TABLE 3.2b CONTINUOUS DATA: GROUP MEANS

Variable	Exptl mean	Control mean	D.F.	t ratio	Prob.
AGE (years)	11.94	11.69	102	4.15	$p < .05$
INDEX OF USE ("All use")	2.06	2.05	89	.08	
ATTITUDES					
Drugs (unspecified)	8.23	8.69	101	1.53	
Alcohol	9.26	9.32	99	.05	
Cigarettes	6.78	6.20	95	1.99	
All Attitudes	24.3	24.2	95	.02	
DRUG KNOWLEDGE					
Drugs (unspecified)	1.04	1.28	89	2.55	
Alcohol	1.14	1.24	99	.46	
Cigarettes	2.41	2.73	95	2.90	
All Drug Knowledge	4.57	5.32	84	5.32	$p < .05$
NON-DRUG KNOWLEDGE					
Parts of the person	3.29	3.56	86	2.37	
Wants vs. Needs	1.39	.98	83	3.05	
Advertising	.71	.80	97	.35	
Decision-making	.82	1.04	99	2.54	
Peer pressure	.82	1.02	100	2.28	
All Non-drug knowledge	7.69	7.67	67	1.37	
ALL KNOWLEDGE	11.15	13.28	60	9.72	$p < .01$

Unspecified drugs were regarded with slight caution (8.23 and 8.69).

Of the twenty-six group comparisons, the only significant difference was seen in pre-intervention levels of drug knowledge. While individual drug knowledge subscales showed no differences, the students in the control classes knew significantly more on the summated drug knowledge index. This difference was responsible for the significant difference on the total, summated All Knowledge index.

No significant differences were seen on the ADES-derived portion of the questionnaire, the nondrug-related knowledge.

The pretest data were also examined using two-way analysis of variance. No main effects for either age or sex were seen in the preprogram data; neither did the corresponding interaction effects (sex by treatment, or age by treatment) generate significant *F*-ratios.

In sum, one or two significant differences out of twenty-five comparisons between experimental and control groups is not greater than one would expect by chance. Students in the control classrooms who submitted matched questionnaires (n=52) can thus be considered equivalent to their counterparts in experimental classes. Finally, of the matched data from Grade Six experimental classes (n=189), a random sample of 52 was selected to provide equal cell sizes. Equal cell sizes were required in the face of (occasional) heterogeneity of variance, as *F* and *t* tests require equal *n*'s when the data showed

heterogeneity of variance (Dixon et al., 1983). Group comparisons using the full sample would have violated the assumptions of the model.

Overall Outcome

Post-test comparisons of experimental and control data suggest some strong and fairly consistent responses in drug-related and (other) curriculum knowlege. Similarly, these comparisons suggest changes in attitudes towards alcohol and tobacco, but--in the ultimately most important measure, behaviour--these analyses suggest little change in reported use of any substance to date (as defined by responses to "Have you ever used .....?").

Table 3.3 shows significant group differences in drug- and nondrug-related knowledge. Tobacco-related knowledge was greater for the experimental students, with a mean 3.45 items answered correctly as opposed to 2.78 for the control group ($t(97)=8.76$, $p < .005$). Similarly, the All Drug Knowledge index, the sum of the three knowledge subscales, showed superior levels for the experimental students ($t(86)=7.55$, $p < .01$). It is also noteworthy that a significant difference in the opposite direction, with control students scoring higher on the All Drug Knowledge index, was evident at pretest.

Regarding nondrug-related knowledge, the program appears to have imparted significant levels of knowledge in all but one of its instructional areas. Albeit using nonvalidated items, Table

Table 3.3 OVERALL OUTCOME: EXPERIMENTAL VS. CONTROL POSTDATA

Table 3.3a CATEGORICAL DATA

No.	Variable	Exptl (n = 52)	Control (n = 52)	D.F.	χ^2 *	Prob.
1	SEX: % Female	48	52	1	.17	
21	CIGARETTE USE: % Yes (more than a puff or two)	16	35	1	3.86	$p < .05$
22	ALCOHOL USE: % Yes (e.g., a bottle of beer, without consent)	10	14	1	.17	
23	CIGARETTE USE: % YES (any, with or without consent)	29	43	1	1.71	
24	BEER USE: % Yes (any, with or without consent)	59	74	1	1.82	
25	WINE USE: % Yes (any, with or without consent)	66	65	1	.00	
26	LIQUOR USE: % Yes (any, with or without consent)	33	43	1	.60	
27	BOGUS USE: % Yes	0	6	N.A.[1]	N.A.[1]	
28	OTHER USE: % Yes (if specified)	4	2	N.A.[1]	N.A.[1]	

*Yates corrected χ^2 [1] Minimum estimated expected value less than 5.

Table 3.3b CONTINUOUS DATA

Variable	Exptl mean	Control mean	D.F.	t ratio	Prob.
AGE	11.64	11.73	102	.81	
INDEX OF USE ("All use")	2.18	2.17	91	.01	
ATTITUDES					
Drugs (unspecified)	5.69	6.41	101	1.87	
Alcohol	7.29	7.65	101	.63	
Cigarettes	5.47	6.88	100	8.82	$p < .005$
All Attitudes	18.55	20.94	100	4.32	$p < .05$
DRUG KNOWLEDGE					
Drugs (unspecified)	1.41	1.38	91	.02	
Alcohol	1.60	1.34	91	.02	
Cigarettes	3.45	2.78	97	8.76	$p < .005$
All Drug Knowledge	6.54	5.52	86	7.55	$p < .01$
NON-DRUG KNOWLEDGE					
Parts of the Person	3.88	3.45	80	4.33	$p < .05$
Wants vs. Needs	1.53	1.16	82	2.10	
Advertising	1.48	.98	98	12.38	$p < .001$
Decision-making	1.22	.90	96	6.43	$p < .05$
Peer pressure	1.18	.89	97	5.15	$p < .05$
All Non-drug knowledge	9.54	7.13	66	19.12	$p < .001$
ALL KNOWLEDGE	15.94	12.75	60	16.37	$p < .001$

3.3 shows that four of the five subscales showed significant differences between the groups. Experimental students scored a higher number of items correctly in: Parts of the Person ($t(80)=4.33$, $p<.05$); Advertising ($t(98)=12.38$, $p<.001$); Decision Making ($t(96)=6.43$, $p<.05$); and Peer Pressure ($t(97)=5.15$, $p<.05$). The summed index of all nondrug knowledge also showed superior scores on the part of the experimental students (9.54 vs. 7.13, $t(66)=19.12$, $p<.001$). Summing the significantly different drug and nondrug knowledge indices produced a significant difference in the All Knowledge index ($t(60)=16.37$, $p<.001$).

The program can also be credited with some change in drug-related attitudes. Beliefs about the use of tobacco--already viewed with equal caution at pretest--showed a further drop for the experimental students, resulting in a significant postprogram difference ($t(100)=11.08$, $p< .005$). This was enough to register a significant difference for the overall attitude index, the sum of the three subscales, even though the other attitude subscales showed no group differences at posttest.

Some change in use of tobacco was reported in the postprogram data. The significant difference in postprogram responses to the "cigarettes: more than a puff or two" item suggests--superficially, at least--success at the behavioural level: 16% of experimental vs. 35% of control students endorsed the item (Chi Square=3.86, $p<.05$). Comparing pre- to post program proportions, though, the experimental rate was less than

at pretest: eight weeks earlier, it was 27% ! Since one's total accumulated use can not be less at posttest, this outcome cannot be totally valid.

The balance of the behavioural responses (items 22 to 26) did not show significantly different proportions of experimental and control groups at posttest. With only one, rather dubious, significant difference in eight group comparisons, these analyses suggest no overall behavioural impact.

Outcome by Sex

Dividing the data into male and female subsets, parallel group comparisons suggest somewhat greater impact is reported by girls than boys. Table 3.4 shows that significantly fewer girls in the control group said they had tried <u>any</u> tobacco (15% vs. 48%, Chi Square=5.11, $p<.05$), as indicated by item 23. The other item measuring cigarette use, however, did not show a significant difference. There was a marked trend across most substance categories for less students of both sexes to report alcohol and tobacco use if they had received the program. For girls, three of the four remaining behavioural items showed group proportions at least 15% different[1].

The tobacco-related portions of the program showed consistent impact for girls. The only attitude subscale to show a significant difference was that regarding tobacco, and it did

---

[1] A cell differential of 25% to 30% was required to reach the .05 level of significance.

so for both sexes. Experimental girls registered more caution than control girls (5.92 vs. 7.19 respectively, $t(42)=4.15$, $p<.05$). Similarly, cigarette knowledge was superior for experimental girls ($t(52)=11.82$, $p<.005$). Tobacco knowledge items were the main contributors to significant difference in the Drug Knowledge index ($t(42)=9.14$, $p<.005$). Finally, two nondrug instructional areas (Parts of the Person and Advertising) showed an advantage for experimental girls ($t(39)=4.22$, $p<.05$ and $t(52)=4.94$, $p<.05$ respectively). The summed drug knowledge index also showed a superior score for the experimental girls.

The impact for boys was not as strong as for girls. Table 3.5 shows no significant differences in behavioural reports between experimental and control boys. Tobacco-related items showed impact in both nonbehavioural traits, while neither the alcohol nor drugs (unspecified) subscales showed significant differences. Specifically, experimental boys knew more about tobacco (subscale means of 3.80 vs. 3.17; $t(46)=6.12$, $p<.05$) and were more cautious about smoking (subscale means of 4.75 vs. 6.54; $t(46)=8.11$, $p<.01$) than control boys. The cigarette attitude differential was substantial enough to register a significant difference in the summated "All Attitudes" index ($t(46)=4.03$, $p<.05$), in spite of virtually identical attitudes towards drugs (unspecified) and small differences in the opposite direction vis-a-vis alcohol attitudes.

Table 3.4 OUTCOME BY SEX: FEMALE EXPERIMENTAL VS. CONTROL POSTDATA

Table 3.4a CATEGORICAL DATA

No.	Variable	Female Exptls	Female Controls	D.F.	χ^2 *	Prob.
21	CIGARETTE USE: % Yes (more than a puff or two)	15	38	1	2.69	
22	ALCOHOL USE: % Yes (e.g., a bottle of beer, without consent)	0	15	1	N.A.[1]	
23	CIGARETTE USE: % YES (any, with or without consent)	15	48	1	5.11	$p < .05$
24	BEER USE: % Yes (any, with or without consent)	48	77	1	3.52	
25	WINE USE: % Yes (any, with or without consent)	46	64	1	1.00	
26	LIQUOR USE: % Yes (any, with or without consent)	28	30	1	.00	
27	BOGUS USE: % Yes	0	4	1	N.A.[1]	
28	OTHER USE: % Yes (if specified)	0	0	1	N.A.[1]	

*Yates corrected χ^2 [1]Minimum estimated expected value less than 5.

Table 3.4b CONTINUOUS DATA

Variable	Female Exptl.	Female Controls	D.F.	t ratio	Prob.
AGE (years)	11.52	11.82	46	4.81	$p < .05$
INDEX OF USE ("All use")	2.00	2.17	N.A.[2]	N.A.[2]	
ATTITUDES					
Drugs (unspecified)	5.19	6.93	52	5.35	$p < .05$
Alcohol	6.85	7.48	52	1.01	
Cigarettes	5.69	7.18	51	5.35	$p < .05$
All Attitudes	17.46	21.59	52	6.09	$p < .05$
DRUG KNOWLEDGE					
Drugs (unspecified)	1.40	1.33	44	.09	
Alcohol	1.50	1.19	50	2.02	
Cigarettes	3.56	2.44	52	11.82	$p < .005$
All Drug Knowledge	6.50	4.83	42	9.14	$p < .005$
NON-DRUG KNOWLEDGE					
Parts of the Person	3.77	3.05	39	4.22	$p < .05$
Wants vs. Needs	1.39	1.18	38	.50	
Advertising	1.37	.89	52	4.94	$p < .05$
Decision-making	.82	.89	51	.13	
Peer pressure	.89	.89	51	.00	
All Non-drug knowledge	8.36	6.62	31	3.38	
ALL KNOWLEDGE	14.8	11.5	26	6.03	$p < .05$

[2]No variance for female exptls.

Table 3.5 OUTCOME BY SEX: MALE EXPERIMENTAL VS. CONTROL POSTDATA

Table 3.5a CATEGORICAL DATA

No.	Variable	Male Exptls	Male Controls	D.F.	χ^2 *	Prob.
1	CIGARETTE USE: % Yes (more than a puff or two)	15	30	1	.84	
22	ALCOHOL USE: % Yes (e.g., a bottle of beer, without consent)	19	14	N.A.[1]	N.A.[1]	
23	CIGARETTE USE: % YES (any, with or without consent)	31	37	1	.04	
24	BEER USE: % Yes (any, with or without consent)	56	71	1	.61	
25	WINE USE: % Yes (any, with or without consent)	46	67	1	1.38	
26	LIQUOR USE: % Yes (any, with or without consent)	25	54	1	3.14	
27	BOGUS USE: % Yes	4	8	N.A.[1]	N.A.[1]	
28	OTHER USE: % Yes (if specified)	0	4	N.A.[1]	N.A.[1]	

*Yates corrected χ^2 [1]Minimum estimated expected value less than 5.

Table 3.5b CONTINUOUS DATA

Variable	Male Exptls	Male Controls	D.F.	t ratio	Prob.
AGE (years)	11.73	11.64	49	.30	
INDEX OF USE ("All use")	2.04	2.17	47	1.37	
ATTITUDES					
Drugs (unspecified)	5.92	5.83	48	.02	
Alcohol	6.73	7.83	48	2.62	
Cigarettes	5.54	6.54	48	2.26	
All Attitudes	18.19	20.21	46	1.46	
DRUG KNOWLEDGE					
Drugs (unspecified)	1.40	1.36	41	.03	
Alcohol	1.23	1.50	48	1.63	
Cigarettes	3.80	3.17	46	6.12	$p < .05$
All Drug Knowledge	6.45	6.14	40	.45	
NON-DRUG KNOWLEDGE					
Parts of the Person	3.76	3.81	40	.04	
Wants vs. Needs	1.17	1.15	41	.01	
Advertising	1.35	1.09	47	1.82	
Decision-making	1.04	.91	45	.52	
Peer pressure	1.20	.87	46	4.77	$p < .05$
All Non-drug knowledge	8.32	7.50	35	1.45	
ALL KNOWLEDGE	14.3	13.5	31	.71	

Program Outcome: Pre vs. Post Experimental Data

CAse Selection

All participating students in the Grade Six classes which received the program were included for the analysis here. Sample size approximated 220 cases for the pre- and post-program questionnaire administrations.

Overall Outcome

Using only the experimental data, pre-post comparisons suggest stronger program impact in knowledge and attitude realm than the previous experimental-control posttest comparisons. Like the previous analyses, no behavioural impact was apparent from these pre-post comparisons.

Specifically, experimental students knew more about cigarettes, alcohol and drugs (unspecified) upon completing the program, compared to their preprogram levels. Table 3.6 shows that the differences in the three drug-related knowledge score are all highly significant ($p<.001$).

Nondrug-related knowledge comparisons were highly significant in three of five areas: Advertising, Decision Making, and Peer Pressure. No significant differentials were apparent for the Parts of the Person or the Wants vs. Needs lessons. Both the summary index for this area (Nondrug Knowledge) and the overall knowledge index (All Knowledge) were

Table 3.6 OVERALL PROGRAM OUTCOME; PRE VS. POST EXPTL DATA

Table 3.6a CATEGORICAL DATA

No.	Variable	EXPTL. PRETEST	EXPTL. POSTTEST	D.F.	χ^2 *	Prob.
1	CIGARETTE USE: % Yes (more than a puff or two)	18	16	1	.04	
22	ALCOHOL USE: % Yes (e.g., a bottle of beer, without consent)	13	15	1	.26	
23	CIGARETTE USE: % YES (any, with or without consent)	25	30	1	.82	
24	BEER USE: % Yes (any, with or without consent)	50	58	1	2.64	
25	WINE USE: % Yes (any, with or without consent)	48	57	1	3.06	
26	LIQUOR USE: % Yes (any, with or without consent)	27	35	1	2.51	
27	BOGUS USE: % Yes	0	1	1	N.A.	
28	OTHER USE: % Yes (if specified)	1	2	1	N.A.	

*Yates corrected χ^2 [1]Minimum estimated expected value less than 5.

Table 3.6b CONTINUOUS DATA

Variable	EXPTL. PRETEST	EXPTL. POSTTEST	D.F.	t ratio	Prob.
INDEX OF USE ("All use")	2.05	2.06	318	.11	
ATTITUDES					
Drugs (unspecified)	5.93	5.46	435	5.06	.025
Alcohol	7.64	7.02	433	9.11	.005
Cigarettes	6.31	5.77	426	9.90	.005
All Attitudes	19.85	18.15	418	14.39	.001
DRUG KNOWLEDGE					
Drugs (unspecified)	1.00	1.35	415	101.09	.001
Alcohol	1.06	1.39	429	21.29	.001
Cigarettes	2.52	3.47	400	23.56	.001
All Drug Knowledge	4.55	6.27	372	126.61	.001
NON-DRUG KNOWLEDGE					
Parts of the Person	3.60	3.78	363	3.41	
Wants vs. Needs	1.27	1.46	382	2.30	
Advertising	.76	1.35	430	71.71	.001
Decision-making	.87	1.09	422	11.23	.001
Peer pressure	.82	1.13	429	21.76	.001
All Non-drug knowledge	7.53	8.88	315	24.05	.001
ALL KNOWLEDGE	12.14	15.10	269	61.93	.001

highly significant ($p<.001$).

All attitude indicators were highly significant. The experimental students reported more caution towards each of the tobacco, alcohol and drugs (unspecified) subscales, as well as the summated All Attitudes index.

As with the earlier experimental-control comparisons, no overall behavioural differentials were significant in these pre-post comparisons.

Outcome by Sex

Once again, the data were divided into male and female subsets. This time, however, boys and girls reported approximately similar program impact. Table 3.8 shows that boys reported more cautious attitudes on all three subscales, and on five of eight knowledge subscales after the program. Girls, on the other hand, reported more cautious attitudes on two of three attitude subscales, and on seven of eight knowledge subscales (shown in Table 3.7). Summated drug-related knowledge and nondrug-related knowledge indices, and the summated attitude measure, were significant for both sexes.

Table 3.7 OUTCOME BY SEX: PRE VS. POST EXPERIMENTAL GIRLS

Table 3.7a CATEGORICAL DATA

No.	Variable	Expt1 Girls Pretest	Expt1 Girls Posttest	D.F.	χ^2 *	Prob.
1	CIGARETTE USE: % Yes (more than a puff or two)	17	16	1	.01	
22	ALCOHOL USE: % Yes (e.g., a bottle of beer, without consent)	11	10	1	.01	
23	CIGARETTE USE: % YES (any, with or without consent)	24	24	1	.00	
24	BEER USE: % Yes (any, with or without consent)	52	57	1	.34	
25	WINE USE: % Yes (any, with or without consent)	45	56	1	2.40	
26	LIQUOR USE: % Yes (any, with or without consent)	28	33	1	.28	
27	BOGUS USE: % Yes	0	0	1	N.A.	
28	OTHER USE: % Yes (if specified)	3	2	1	N.A.	

*Yates corrected χ^2 [1]Minimum estimated expected value less than 5.

Table 3.7b CONTINUOUS DATA

Variable	Expt1 Girls Pretest	Expt1 Girls Posttest	D.F.	t ratio	Prob.
INDEX OF USE ("All use")	2.09	2.07	172	.05	
ATTITUDES					
Drugs (unspecified)	5.86	5.54	225	1.33	
Alcohol	7.56	7.00	226	4.30	.05
Cigarettes	6.29	5.73	220	4.70	.05
All Attitudes	19.71	18.21	216	6.07	.05
DRUG KNOWLEDGE					
Drugs (unspecified)	2.48	3.33	220	40.96	.001
Alcohol	1.03	1.50	221	21.84	.001
Cigarettes	.99	1.33	216	12.00	.001
All Drug Knowledge	4.51	6.22	204	66.06	.001
NON-DRUG KNOWLEDGE					
Parts of the Person	3.65	3.83	195	1.72	
Wants vs. Needs	1.30	1.68	195	5.32	.05
Advertising	.81	1.32	226	27.92	.001
Decision-making	.93	1.12	221	4.02	.05
Peer pressure	.85	1.19	222	14.45	.001
All Non-drug knowledge	7.97	9.13	165	8.37	.005
ALL KNOWLEDGE	12.47	15.40	149	31.75	.001

Table 3.8 OUTCOME BY SEX: PRE VS. POST EXPTL, BOYS

Table 3.8a CATEGORICAL DATA

No.	Variable	Expt1 Boys Pretest	Expti Boys Posttest	D.F.	χ^2 *	Prob.
1	CIGARETTE USE: % Yes (more than a puff or two)	18	17	1	.01	
22	ALCOHOL USE: % Yes (e.g., a bottle of beer, without consent)	15	20	1	.75	
23	CIGARETTE USE: % YES (any, with or without consent)	26	36	1	1.69	
24	BEER USE: % Yes (any, with or without consent)	46	59	1	3.12	
25	WINE USE: % Yes (any, with or without consent)	51	58	1	.79	
26	LIQUOR USE: % Yes (any, with or without consent)	23	37	1	3.38	
27	BOGUS USE: % Yes	0	2	1	N.A.*	
28	OTHER USE: % Yes (if specified)	0	2	1	N.A.*	

*Yates corrected χ^2 [1]Minimum estimated expected value less than 5.

Table 3.8b CONTINUOUS DATA

Variable	Expt1 Boys Pretest	Expti Boys Posttest	D.F.	t ratio	Prob.
INDEX OF USE ("All use")	2.00	2.06	191	N.A.**	
ATTITUDES					
Drugs (unspecified)	6.01	5.38	206	4.14	.05
Alcohol	7.71	7.05	203	4.43	.05
Cigarettes	6.30	5.62	202	4.64	.05
All Attitudes	19.96	18.09	198	7.77	.01
DRUG KNOWLEDGE					
Drugs (unspecified)	2.57	3.62	193	63.28	.001
Alcohol	1.09	1.27	206	3.23	
Cigarettes	1.00	1.36	182	11.42	.001
All Drug Knowledge	4.61	6.32	166	59.89	.001
NON-DRUG KNOWLEDGE					
Parts of the Person	3.53	3.72	164	1.90	
Wants vs. Needs	1.24	1.22	1.83	.02	
Advertising	.71	1.37	203	45.90	.001
Decision-making	.80	1.06	199	8.23	.005
Peer pressure	.79	1.07	205	8.11	.005
All Non-drug knowledge	7.08	8.57	148	16.89	.001
ALL KNOWLEDGE	11.73	14.72	118	30.46	.001

** - No varaince among pretest data.

Results: Comparing the Two Analytical Views

A simultaneous look at the two sets of results is useful. Attitude and knowledge measures from Tables 3 through 8 have been combined in Table 3.9 to show a side-by-side presentation of the experimental vs. control post-data, and the pre vs. post experimental data analyses. The behavioral measures are not repeated in Table 3.9 because they were interpreted as showing no effect earlier in this chapter.

Behaviour

The behavioral indicators show no significant differences in either set of analyses, with two exceptions. Table 3.3 showed a significant difference between experimental (16%) and control (33%) students on item 23. Upon closer inspection, though, the former figure had dropped from the earlier pretest figure (27%, shown in Table 3.2), rendering their lifetime index without meaning. This and all other experimental-control comparisons, showed the same lack of effect overall as did Table 3.6, the pre- vs. post-program experimental data.

The other exception to the complete lack of significant differences in the behavioural measures appeared in Table 3.4. On item 23, cigarette use, girls who received the lessons reported less use than girls in the control group at posttest. Item 21, also comparing cigarette use at posttest, did not show a significant difference. Moreover, neither item in the pre-post comparisons (table 3.7) showed a significant difference. All

Table 3.9 COMPARISON OF RESULTS FROM THE TWO SETS OF ANALYSES

Significance Levels from Experimental-Control Postprogram Comparisons vs. Pre-Post Experimental Comparisons: Attitude and Knowledge Measures

Variable	Experimental-Control Postprogram Comparisons		
	Overall Outcome (from Table 3.3)	Outcome for Girls (from Table 3.4)	Outcome For Boys (from Table 3.5)
ATTITUDES			
Drugs (unspecified)		.05	
Alcohol			
Cigarettes	.005	.05	
All Attitudes	.05	.05	
DRUG KNOWLEDGE			
Drugs (unspecified)			
Alcohol			
Cigarettes	.005	.005	.05
All Drug Knowledge	.01	.005	
NON-DRUG KNOWLEDGE			
Parts of the Person	.05	.05	
Wants vs. Needs			
Advertising	.001	.05	
Decision-making	.05		
Peer pressure	.05		.05
All Non-drug knowledge	.001		
ALL KNOWLEDGE	.001	.05	

Variable	Pre-Post Experimental Comparisons		
	Overall Outcome (from Table 3.6)	Outcome for Girls (from Table 3.7)	Outcome for Boys (from Table 3.8)
ATTITUDES			
Drugs (unspecified)	.05		.05
Alcohol	.005	.05	.05
Cigarettes	.005	.05	.05
All Attitudes	.001	.05	.01
DRUG KNOWLEDGE			
Drugs (unspecified)	.001	.001	.001
Alcohol	.001	.001	
Cigarettes	.001	.001	.001
All Drug Knowledge	.001	.001	.001
NON-DRUG KNOWLEDGE			
Parts of the Person			
Wants vs. Needs		.05	
Advertising	.001	.001	.001
Decision-making	.001	.05	.005
Peer pressure	.001	.001	.005
All Non-drug knowledge	.001	.005	.001
ALL KNOWLEDGE	.001	.001	.001

tolled, then, three of four comparisons suggest no behavioural impact for girls, while the fourth is significant.

Knowledge

Among the knowledge measures, the two sets of analyses suggest somewhat different impact. The pre-post differences among the experimental group suggest much more comprehensive impact than the experimental-control differences, especially when the data for boys and girls are examined separately.

Overall, all four of the pre-post comparisons using the drug knowledge scores were significantly different, while only two (Cigarette and Drug Knowledge) were significantly different among the experimental-control comparisons. Among nondrug knowledge scores, the strength of program impact appears about the same. Among both sets of analyses, though, the summated Drug, Nondrug and All Knowledge indices comparisons were significant--suggesting equally effective knowledge changes overall.

The task of interpreting the knowledge outcome for girls is more difficult, however, inasmuch as the two analyses suggest somewhat different outcome. Only two of four drug knowledge scores were significant for girls in the experimental-control comparisons, while all four were significant in the pre-post comparisons. Similarly, the nondrug knowledge areas show stronger effect according to the pre-post comparisons: 5 of 6 areas, compared to 2 of 6 for the experimental-control analyses,

were significantly different. The discordance may be due to sample size: degrees of freedom were about 200 for the former and only about 50 for the latter. The summated All Knowledge index, however, showed a significant impact for girls in both sets of analyses.

A dramatic difference in outcome is apparent in the boys' knowledge scores. Among the drug knowledge subscales, only one experimental-control comparison achieved significance, while 3 of 4 pre-post comparisons were different. Likewise in the nondrug knowledge area, one of six compared to four of six comparisons (respectively) were significantly different. Again, the difference in sugggested outcome between the two sets of analyses may be a reflection of the four-to-one disparity in sample size.

Girls came out ahead in the knowledge scores, whatever set of analyses was used. For the experimental-control comparisons, girls showed five of 11 significant differences in subscales and summated indices, contrasted to two for boys. Likewise, the pre-post analyses show girls with 10 of 11 significant differences, compared with eight for boys. Whichever set of analyses are used, girls came out ahead in the knowledge realm.

Attitudes

Two somewhat different views of outcome vis-a-vis attitude are apparent, with the pre-post comparisons again suggesting stronger impact overall. Specifically, the Alcohol and the Drugs

(Unspecified) Attitude subscales were significant only in the pre-post comparisons. Both sets of analyses, though, suggested significant impact in the Cigarette Attitude subscale and in the summated All Attitude index.

The outcome for boys was also more favorable acoording to the pre-post comparisons. Their attitudes were significantly different in three of the four pre-post comparisons, but in only one experimental-control comparison.

The degree of change in girls' attitudes towards the three drug categories appeared to be about the same for both sets of analyses. Three of four comparisons were significantly different in each set of analyses.

Analysis of Covariance

Fifteen analyses of covariance were performed on the postdata subscales and indices, using each student's respective pretest score as the covariate. Regression coefficients for experimental and control groups, as well as tests for equality of slopes, were computed.

Eight of fifteen groups comparisons were significantly different. Five were subscales (Cigarette Knowledge and Attitudes, Parts of the Person, Advertising, and Peer Pressure); the remainder were the knowledge indices (Drug, Nondrug and All Knowledge). Tests of underlying assumptions, however, did not allow any of the eight to stand, so the data are not displayed.

The only variables to show statistically similar slopes in the respective experimental and control group comparisons were Parts of the Person and All Knowldege. Three further variables showed substantially similar slopes--Cigarette Knowledge, Drug Knowledge and Peer Pressure. Five of the eight variables, then, passed the "homogeneity of regression" test. Unfortunately, none of the measures showed linearity of regression--the detrended normal probability plots of all five were curvilinear.

Validity

If the questionnaire truly represents student knowledge, it should stand up to scrutiny under item analysis. Similarly, students' reported behaviour, if accurate, should show minimal amounts of under- and over-reporting. Finally, attitude items should correlate well within their subscale, but not between subscales. As it turns out, the questionnaire performed reasonably well on these tests.

Over-reporting

One way of measuring deliberate over-reporting is to include an item asking about students' experience with a phony substance. In this case, use of "aliphatics (stars, VMA)" was not reported on any of the matched Grade Six questionnaires at pretest. It was endorsed by three control students and none of the experimental group at posttest.

Under-reporting: Inconsistent Responses

If students were deliberately under-reporting, we would expect some of that under-estimation to be evident in their reporting less total accumulated use of some substances at posttest than at pretest eight weeks earlier.

Since the questionnaire includes some relatively fine gradations of drug use, the calculations allow for slight variation (e.g. if a student interpreted consumption of four sips of wine on two occasions as answer "e" at pretest and "d" at posttest). Table 3.10 gives proportions of students who

Table 3.10 Inconsistent responses to Behavioural Items

Substance	Sample Size	% reporting less use on 2nd questionnaire
Cigarettes	216	1.9
Beer	218	0.9
Wine	214	1.4
Liquor	198	1.5
"Aliphatics"	206	-
Other (marijuana)	195	-
Other (misc.)	241	0.8

reported postprogram use at least two levels of use lower than pretest.

The percentages of inconsistent responses for all of the different substances are between 0.8% and 1.9%. The proportion of postprogram responses that were only one category less than at pretest ranged from 0.8% to 9.6% (not shown).

Knowledge Items: Percentage Gains

The items in this questionnaire were chosen because they had been used in other research. Some items or groups of items have been subjected to validity checks, others have not (see Instrumentation, p. 32). As a result, effects of this year's instruction may be underrated, i.e. nondiscriminating items may contribute to error variance in some subscales and indices. Type II errors may also result from a failure to statistically adjust for prescores in postprogram group comparisons (see Results: ANCOVA). On the other hand, nondiscriminating items may reflect less-than-desired teaching efficacy. Only further research can distinguish between the two hypotheses.

Table 3.11 compares the proportions of correct answers for knowedge items at pre- and posttest. Ten of the twenty-seven knowledge items, including five of six which comprise the "Wants vs. Needs" subscale (items 34-39), failed to discriminate by more than 5% (indicated by "X" in the table). A further five items showed a gain of less than 10% in correct answers ("?" in table). None of the "Parts of the Person" (items 29 - 33) series

Table 3.11 Item Analysis: Proportion of exptl. students endorsing correct answer at pre- and post-test

Item No.	% correct at pre-test	% correct at post-test	Gain	Item No.	% correct at pre-test	% correct at post-test	Gain
5	79	85	?	19	25	33	?
6	48	88		20	59	79	
7	28	43		29	76	84	?
8	51	72		30	65	73	?
9	10	14	x	31	92	95	x
10	65	82		32	25	31	?
11	11	39		33	96	92	x
13	51	65		34	9	14	x
14	94	93	x	35	32	48	
15	21	47		36	27	30	x
16	27	51		37	19	22	x
17	53	85		38	20	21	x
18	63	74		39	15	12	x

***Data from all experimental students (n=189).**

Legend: "X" - indicates item that failed to discriminate by more than 5%
"?" - indicates a gain of less than 10%

discriminated well (more than 10%). These two instructional areas showed similar measurement and/or instructional difficulties in the ADES (1981) report.

Attitude Correlations

All attitude items collected from Grade Six, matched questionnaires correlated well with their own subscale, and poorly with the items and total scores for other subscales. Table 3.12 shows that items pertaining to attitudes about alcohol correlated with the total alcohol subscale score reasonably well (.48 to .65). Cigarette attitude items correlated among themselves somewhat better (.60 to .73). The drugs (unspecified) items all showed fair to good correlations (.43 to .66).

The correlation matrix also supports claims for discriminant validity in the attitude portion of the questionnaire: no "cross correlations" exceeded the .30 range.

The correlations for each individual item with the overall attitude index were generally low, ranging from .24 to .43. The subscales, fortunately, showed consistent agreement with the summated attitude scale (.56 to .59).

Knowledge Correlations

Standardized items in the knowledge realm are rare, since the specific nature of the information taught in the classroom,

Table 3.12 ATTITUDE CORRELATIONS

ITEM/ SUBSCALE	40	41	42	ALCOHOL SUB-SCALE	43	44	45	CIGA-RETTE SUBSCALE	46	47	48	DRUGS (unspec'd) SUBSCALE
40. Alcohol												
41. Alcohol	-.03											
42. Alcohol	.19	-.13										
ALCOHOL SUBSCALE	.65	.48	.62									
43. Cigarettes	-.03	.15	-.01	.06								
44. Cigarettes	-.20	.01	-.01	-.14	.12							
45. Cigarettes	-.18	-.11	.03	-.14	.04	.27						
CIGARETTES SUBSCALE	-.20	.04	-.03	-.10	.60	.73	.65					
46. Drugs (unspec'd)	-.17	.03	-.14	-.16	.00	.15	.22	.18				
47. Drugs (unspec'd)	.30	-.10	.14	.20	-.06	-.09	-.21	-.19	-.22			
48. Drugs (unspec'd)	-.04	.23	-.19	-.01	.11	.13	.12	.17	.22	-.14		
DRUGS (unspec'd) SUBSCALE	.08	.09	.11	.03	.03	.10	.06	.08	.56	.43	.66	
ALL ATTITUDES INDEX	.31	.36	.31	.56	.42	.41	.32	.58	.30	.24	.43	.59

which in turn depends on the program's objectives, makes for much more variability. For example, the definition of "alveoli" may be important for one program's physiological component, but not mentioned in another. Similarly, the identification of "hidden persuaders" in advertisements is an integral part of the decision-making component of another. The knowledge taught by each program is almost unique from an evaluator's point of view, yet both programs teach (and wish an evaluation to show) more cautious or conservative attitudes toward cigarette smoking. Unlike the more general attitude measures, then, some of the items that test specialized knowledge have not been subjected to validation procedures.

The correlation matrix of the knowledge subscales and the (related) overall indices is shown in Table 3.13. Table 3.13a shows moderate correlations involving each of the three drug knowledge subscales with the All Drug Knowledge index (ALLDRKN) (.55 to .65). Scores among individual subscales were not related.

A similar pattern among nondrug knowledge scores is evident in Table 3.13b. Correlations of the subscales with the Nondrug Knowledge index (NONDRKN) range from .46 to .57. With the exception of the association between the Decision Making and Advertising knowledge subscales, which are not independent (one questionnaire item went into both subscales), all subscale intercorrelations are at or below .28.

Table 3.13 KNOWLEDGE CORRELATIONS

Table 3.13a Drug Knowledge

	Cigarettes	Alcohol	Drugs	All Dr. Kn.
Alcohol	.12			
Drugs	.01	.00		
All Dr. Kn.	.65	.55	.55	
All Kn.	.45	.35	.34	.66

Table 3.13b Nondrug Knowledge

	Parts of Person	Wants vs. Needs	Advertising	Decision-making*	Peer Pressure*	Non-Drug Kn.
Wants vs. Needs	.04					
Advertising	.19	.06				
Decision-making*	.12	-.03	.28			
Peer Pressure*	.13	-.03	.28	N/A		
Non-drug Knowledge	.46	.57	.57	.57	.57	
All Knowledge	.36	.50	.60	.48	.48	.90

*Peer pressure and decision-making not independent

The "bottom line" in the two tables is ALLKN--the summed total of all the drug and nondrug knowledge items. Correlations between ALLKN and the three drug knowledge subscales ranged from .35 to .45, with their summed index (ALLDRKN) correlating well with the overall knowledge index (.66). The individual nondrug knowledge subscales fared better than their counterparts, with fair-to-good correlations of .36 to .60. The nondrug knowledge total score correlated highly (.90) with the All Knowledge total score.

IV. DISCUSSION

While the program did not achieve behavioural changes, both sets of analyses suggest results that go beyond the immediate post-program objectives set out by its sponsor. A behavioural program goal was not considered realistic by ADES.

Attitude changes are suggested by both sets of analyses, even though they were not set as a postprogram goal. Students who received the lessons had significantly different attitudes towards cigarettes, compared to both their own previous attitudes and compared to those of students who didn't receive the lessons. The difference was present for both boys and girls, when examined separately. All other attitude subscales and indices except one, Alcohol Attitudes among boys, scored significant differences in one, but not both, of the analytical views. In sum, then, the program can be credited with an immediate change in cigarette attitudes among all students, and there is some evidence for attitude change in all three subscales among boys and girls. This outcome is an unexpected bonus.

What was set out as the program objective--knowledge changes--has been achieved, at least for the most part. For all

students taken as a group, Cigarette Knowledge and All Drug Knowledge comparisons were significant in both sets of analyses. Similarly, three of five nondrug knowledge subscales, as well as the Nondrug Knowledge and the All Knowledge indices, were also significant for both sets of comparisons.

Neither the agency nor the author predicted differential knowledge outcome for the sexes. While knowledge regarding cigarette and advertising appears to have been absorbed equally among boys and girls, the latter gained substantially more knowledge according to the experimental-control comparisons. A smaller advantage is suggested for girls in the pre-post comparisons.

Although the program has performed quite admirably, the interpretation of the results must be seen in light of a number of considerations regarding internal and external validity. Regarding the former, the research design, sample selection, instrumentation, and statistical analysis deserve comment.

Design

Relying on teachers to set and communicate their own timetable was considered necessary, since this was the first year ADES staff had not taught the lessons themselves. In previous years, ADES staffing considerations had justified requests for specific starting dates in both early and late spring. The "waiting list control" design that was initially planned required that the timetables of all twenty or thirty

participating teachers be submitted by early February, in order that those starting in late April could serve as "waiting list control" groups for the classes starting in late February. That expectation was unrealistic. Often the teachers themselves were not able to specify a starting date until April, as they felt a need to complete their own mandatory curriculum before starting supplementary material. Future evaluators will be more able to set a realistic research timeline, since teachers will draw from their previous experience.

Two lessons regarding research design can be drawn from this experience. First, the simplest possible design that is adequate for the job should be used. In this case, the fall-back option of pre/post control group design, as well as the attendant timeline should at least have been specified in advance. Such a plan would have allowed time for a larger number of control groups to be recruited. Equal sample sizes, as well as Grade Seven control groups, would have been the result.

Larger numbers in the control condition would have allowed exclusion of classrooms where unexpected complications arose, viz. unruly reunion of peer groups where two classrooms were combined for the questionnaire administration, and a boisterous class had just returned from a heated game of soccer. Another source of control data in this evaluation was overlooked: in at least two classrooms, predata were collected from classrooms scheduled to take part in the program, but no lessons were actually delivered. A return trip for postdata would have had

the effect of adding those two classes to the control group.

Case Selection

The biggest problem of internal validity in the evaluation is the exclusion of the "non-matched" control data. Had the control groups been more numerous, the (hypothesized) situational complications could have been eliminated by the simple exclusion of the "dirty" classrooms. As it stands, it is not possible to distinguish the source of the "matching" inequality: it could either be intrinsic to the no-treatment control group, e.g. plausibility of the research protocol as genuine, or extrinsic, i.e. the unique situations described above. If it is the former, further evaluators may wish to compare experimental groups with an attention-placebo or a minimum-treatment control group, in addition to or in place of the no-treatment control group used here.

Measures

The questionnare needs to be improved. While the various tests of internal consistency held up well, especially regarding the attitude scores, the instrument shed no light on two nondrug curriculum knowledge areas. Ten of fourteen questions that gained less than 10% between pre- and posttest came from two lessons, "Parts of the Person" and "Wants vs. Needs". These areas also plagued the earlier in-house evaluation done by ADES in 1981. A serious reworking of these items, perhaps other than

paper-and-pencil measures, is needed.

The attitude items, borrowed from Goodstadt et al. (1978) used a Likert-type scale to assess general attitudes toward alcohol, cigarettes and drugs (unspecified). Osgood's Semantic Differential technique (Snider and Osgood, 1969) could be explored in future research, as could the development of items that separately address cognitive, affective and conative components of attitudes.

The index of drug use, developed by Blum et al. (1978) for a wider age range, was not appopriate for this Grade Six cohort. Alcohol and tobacco use had barely begun for the vast majority of the eleven and twelve year old students. In addition, the effect of any narrow age range, old or young, may render the index inoperative.

The "lifetime use" items (23 - 28) were not satisfactory. Endorsement of "c" through "e" were originally intended to reflect total, accumulated use. Endorsement of answers "d" through "f" on each of those items was to measure current use. That distinction was not sufficiently clear, especially inasmuch as the "current use" reponses did not specify a time frame. As a result, these behavioural responses have been interpreted of lifetime use--not a particularly sensitive level of measurement. Trying to combine both current and lifetime use in one item, then, resulted in getting neither. ******* Is this fair? *******

In the present research, attitudes, knowledge and behaviour have been treated as independent measures of program impact.

Future investigations would do well to explore the extent of their interrelationships, especially between the conative attitude (often referred to as behavioural intentions) and the behaviours themselves. Those data would go a long way towards resolving the debate vis-a-vis immediate program goals and ultimate behavioural objectives.

Finally, the ghost of the multitrait-multimethod prescription (Campbell and Fiske, 1959) continues to haunt this and most other DREP evaluations for use of only the ubiquitious questionnaire. Physiological methods could be borrowed from the Smoking and Health Education Program (SHEP) evaluation literature: biochemical assays of saliva samples as a test for cigarette smoking (Benfari et al, 1977; Hurd et al., 1980; Prue et al., 1981) have proved useful as a supplementary method for gathering behavioural outcome data, at least in older cohorts. While I saw no reports in the literature of physiological measures with drug education programs, such exploratory efforts may be worthwhile. School records have also been used to circumvent the exclusive reliance on questionnaires, although the reluctance of school administrations may account for the rarity of this source of data.

Limitations

Apart from the methodological comments affecting internal validity, four aspects of the present evaluation limit the generalizability of the results. First, data were collected from Grade Six classes only: evaluation of the program as a whole can only occur with the inclusion of Grade Seven students.

Second, immediate program impact does not imply durable impact. To have begun the evaluation process with an assessment of effects after six or twelve months would have been putting the cart before the horse. A two month follow-up component was originally proposed for this research, but was not possible because the program was offered so late in the school year. Follow-up after any longer period of time was beyond the scope of this project.

Third, the assumption shared by ADES and much of the DREP literature is that limiting and/or delaying the onset of any drug use in pre- or early adolescent years will result in healthier, happier people in later adolescence or adulthood. Such an assumption affects the choice of measures, depending on age level of the target population. For a Grade Six program, it seems appropriate at posttest and even at one or two year follow-up to measure overall use of tobacco and alcohol, since the distinction between "use" and "abuse" may not exist for the vast majority of twelve year olds. On the other hand, where sixteen or eighteen year olds are the subjects of an evaluation--either as part of a long term follow-up component to

an ADES-type program, or in a program for senior high shcool students--evaluators may want to measure problematic as well as overall consumption of tobacco, alcohol, cannabis and other illicit substances.

Finally, while the program's sponsors will no doubt be pleased to see that their "Making Decisions" curriculum works, it remains to determine exactly what it is about the eight lessons that "works". Further evaluation of this program is needed to determine, for example, the effect of experienced vs. inexperienced instructors, whether classroom teachers or outside experts. Comparing this with other DREP's, either in a future evaluation project or awaiting publication of results from other well evaluated short- to medium length programs, would shed further light on differential outcome as per alcohol vs. tobacco vs. cannabis with the Grade Six students. Similarly, a "sleeper" effect may occur with programs using the affective, but not for the didactic strategy.

Finally, even if the most effective components from "Making Decisions" and other programs are combined, evaluators and program sponsors must not be saddled with unrealistic expectations, as they were during the "drug explosion" of the late 1960s and early 1970s. Even if the fruits of the best efforts are combined, how happy will administrators and politicians be if the lessons account for five per cent of the variance?

APPENDIX I

PARENT CONSENT FORM

As part of our Health Curriculum this year, our Grade Six classes are planning to participate in a Decision-Making Program with the nonmedical use of drugs as the central theme. This program, entitled "MAKING DECISIONS: An Approach to Prevention", has been developed by the Alcohol-Drug Education Service (ADES) of Vancouver, under a grant from the B.C. Ministry of Health. It has been tested in numerous Lower Mainland schools.

We will be holding a meeting with parents to discuss the program at __(time)__ on __(date)__, at the school. Mr. Colin Mangham of ADES will explain the purposes of the program and answer questions from parents at that time.

"MAKING DECISIONS" is an eight-lesson course in which are discussed factors in decision-making such as peer pressure, advertising, and the difference between needs and wants. The course gives information on alcohol and tobacco, and discusses alternatives to nonmedical drug use. A pre/post/follow-up survey will be used to measure the effectiveness of the program.

Please return the consent form below to the school prior to __(date)__.

Thank you,

(principal)

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

CONSENT FORM

Name(s) of child:
I would (not) like my child to take part in the described program.

I will (not) be attending the parent night on _____.

(Signature of parent or guardian)

APPENDIX II

STUDENT INFORMATION SHEET AND QUESTIONNAIRE

You are being asked to take part in a project that tries to measure the value of a series of lessons about decision making. The lessons will discuss concern for others, the difference between needs and wants, advertising, the non-medical use of drugs, etc. We will ask you to fill out three questionnaires -- one now, one after you receive the program, and the third two months after the end of the program.

In this questionnaire we'd like you to be honest about recording your answers, since all information from students is important for this research. Since you are <u>not</u> to put your name on the questionnaire, no one but you will know who completed it. All answers will be seen only by the research staff. Remember, your participation is completely voluntary.

To allow us to analyze changes before and after the program, we would like you to make up a secret identity code. This procedure enables us to identify each <u>set</u> of three questionnaires, without knowing the identity of any student. Your secret code, to be written in the space provided below, comes from three places:

a) The first part of the secret code comes from the last letter in your family name.
b) The second part of the scret code comes from the last digit in your telephone number.
c) The third part of the secret code comes from the last digit of your street address.

EXAMPLE: John Chang 555-2894 1052 East 23rd Avenue.

John's Secret Code g 4 2

YOUR SECRET CODE: ____ ____ ____

1. Sex: a) female b) male

2. Age: a) 10 or under b) 11 c) 12 d) 13 or older

3. Grade in school: a) 5 b) 6 c) 7

Read each of the statements listed below. Circle:

The letter 'A' if it is TRUE.

The letter 'B' if it is FALSE.

The letter 'C' if you are NOT SURE.

		TRUE	FALSE	NOT SURE
5.	Smokers are more likely than non-smokers to develop lung cancer.	A	B	C
6.	Nicotine is the drug in cigarettes.	A	B	C
7.	Alcohol is a depressant.	A	B	C
8.	Addicting means habit-forming.	A	B	C
9.	Alcohol is digested like food.	A	B	C
10.	Alcohol is a drug.	A	B	C
11.	Nicotine is a stimulant.	A	B	C
12.	Although we may do much thinking in a day, we make few decisions.	A	B	C

13. Melissa and Randy have been asked by their teacher to bring three containers, all of which contain drugs which are legal, but which are not medicines. They should bring:

a) cough syrup, Aspirin, and Tylenol containers

b) Aspirin, antibiotic, and vitamin containers

c) sugar, vitamin, and soft drink containers

d) beer, cigarette, and coffee containers

e) none of the above.

14. Mr. John Smith has smoked cigarettes for a long time, and seems healthy. He says that this shows that cigarettes don't affect him, and that it would be silly for him to quit for reasons of health. Do you agree with him?

a) yes b) no c) don't know

15. The ingredient in cigarette smoke that is believed to be the cause of lung cancer is:

a) nicotine
b) carbon monoxide
c) tobacco tars
d) charcoal

.../2

16. Through advertising, people are persuaded to buy certain brands of cigarettes because they:
    a) are under the subtle influence of suggestion.
    b) identify cigarette use with social pleasures.
    c) connect smoking with beautiful, sophisticated people.
    d) are influenced by all of the above.
    e) are influenced by none of the above.
    f) not sure.

17. Bob's Fast Foods ran an advertisement in a local magazine saying: "Join the Crowd, Eat at Bob's."
    Bob:
    a) is spending a lot of money.
    b) is smart.
    c) has very good food.
    d) has a good restaurant.
    e) is using a "hidden persuader."
    f) not sure

18. If you were a "thoughtful decider", you would probably make many decisions in the following way:
    a) Look at the problem. Decide. Then think carefully about the results.
    b) Look at the problem. Think carefully about the choices, then decide.
    c) Decide quickly, then think carefully about the results.
    d) Look carefully at the problem, then ask someone older to decide for you.
    e) Not sure.

19. Pretend that this evening you have planned to finish your science project which is due tomorrow. Two friends stop by at suppertime and ask you to go to a rollerskating party with them. Your parents say you may go if you wish. As a "thoughtful decider", you would think about which of the following before you decide?
    a) the difference between needs and wants.
    b) peer pressure.
    c) what choices you have (what will happen if you go or if you don't go).
    d) all of the above.
    e) not sure.

20. Tanya's friends want her to stay and play, even though she has homework to do. She decides to stay. Tanya is most probably being affected by:
    a) advertising.
    b) her choices.
    c) group pressure.
    d) her parents.
    e) none of the above.

21. Have you ever smoked cigarettes? If so, when was the first time that you took more than just a puff or two?
    a) No
    b) before age 9.
    c) at age 9.
    d) age 10.
    e) age 11.
    f) age 12.
    g) age 13 or older.

22. Have you ever drank an alcoholic beverage (for example, a bottle of beer) without your parents' consent? If so, when was the first time?
    a) No.
    b) before age 9.
    c) at age 9.
    d) age 10.
    e) age 11.
    f) age 12.
    g) age 13 or older.

Circle the answer that best describes your use, if any, of the following substances (without your parents' permission).

FREQUENCY	23. Cigarettes	24. Alcoholic Beverages: Beer	25. Alcoholic Beverages: Wine	26. Alcoholic Beverages: Liquor	27. Aliphatics	28. Other (please specify)
Never tried it.	A	A	A	A	A	A
Tried it once or twice.	B	B	B	B	B	B
Tried it several times (but less than 10).	C	C	C	C	C	C
Less than once per week (but more than 10 times).	D	D	D	D	D	D
Two or three times/week.	E	E	E	E	E	E
Daily or almost daily.	F	F	F	F	F	F

It can be said everybody has four parts to their person. These are BODY, FEELINGS, MIND, and BELIEFS AND VALUES. What part of you is most involved in each of the following activities?

ACTIVITIES	BODY	FEELINGS (Emotions)	MIND	VALUES AND BELIEFS
29. Doing homework three days ahead of time.	____	____	____	____
30. Taking part in religious services.	____	____	____	____
31. Doing push-ups.	____	____	____	____
32. Returning a five dollar bill to a person who just dropped it (but didn't realize it).	____	____	____	____
33. Saying goodbye to your best friend who is moving away.	____	____	____	____

Sometimes the words "want" and "need" are used to mean the same thing; at other times they have very different meanings. In the following set of questions, circle:

"1" if the item is always a "want".

"2" if the item is usually a "want".

"3" if it doesn't matter which word is used.

"4" if the item is usually a "need".

"5" if the item is always a "need".

EXAMPLE: FOOD Want 1 2 3 4 (5) Need

		Always Use "Want"	Usually Use "Want"	Doesn't Matter	Usually Use "Need"	Always Use "Need"
34.	Skateboard	1	2	3	4	5
35.	Love	1	2	3	4	5
36.	TV	1	2	3	4	5
37.	Bicycle	1	2	3	4	5
38.	Friends	1	2	3	4	5
39.	Clothes	1	2	3	4	5

.../4

Mark an "X" in the spot that indicates your feelings about each statement below.

EXAMPLE: Christmas holidays should be longer.

Strongly Agree	Agree	Don't Know	Disagree	Strongly Disagree
X				

		Strongly Agree	Agree	Don't Know	Disagree	Strongly Disagree
40.	Alcohol is a good thing to have at a party because it helps people have more fun.					
41.	The dangers of alcohol outweigh the pleasures of drinking.					
42.	Too much fuss is made about alcohol use.					
43.	Most cigarette advertising is misleading.					
44.	The sale of tobacco (cigarettes) should be banned.					
45.	Adults should set a good example by not smoking (tobacco).					
46.	We need stricter control of drugs.					
47.	There is nothing wrong with using drugs if they make you feel good.					
48.	Something is wrong with the world when regular drug taking becomes an accepted way of life.					
49.	All drugs are harmful.					

REFERENCES

Alcohol-Drug Education Services (ADES). (1981, September). Report on Evaluation of a Grade Six Program: "Making Decisions: An Approach to Prevention" 1980-81. Unpublished manuscript. Vancouver, B.C.,

Amendolara, F.R. (1973). Modifying attitudes towards drugs in seventh grade students. Journal of Drug Education, 3, 71-78.

Baker, R.J. (1973). Drug education: Is it doing any good? Education Digest, 38, 38-40.

Bennett, C.A. and Lumsdaine, A.A. (1975). Evaluation and Experiment New York: Academic Press.

Berberian, R., Gross, C., Lovejoy, J. and Paparella, S. (1976). The effectiveness of drug education programs: A critical review. Health Education Monographs, 4, 327-76.

Berberian, R., Thompson, W., Kasl, S., Gould, L. and Kleber, H. (1976). The relationship between drug education in New Haven schools and changes in drug use and drug-related beliefs and perceptions. Health Education Monographs, 4, 327-76.

Blum, R.H. (1976). Drug Education: Results and Recommendations Lexington, Mass.: Lexington Books.

Blum, R.H., Garfield, Emily, Johnstone, Judy, and Magistad, J.G. (1978). Drug education: Further results and recommendations. Journal of Drug Issues, 8, 379-426.

Braucht, G.N., Follingstad, Diane, Brakarsh, D. and Berry, K.L. (1973). Drug education: A review of goals, approaches, effectiveness, and a paradigm for evaluation. Quarterly Journal of Studies on Alcohol, 34, 1279-92.

Braucht, G.N., Brakarsh, D., Follingstad, Diane, and Berry, K.L. (1973). Devient drug use in adolescence. Psychological Bulletin, 79, 92-106.

Campbell, Donald and Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod Matrix. Psychological Bulletin, 56, 81-105.

Campbell, Donald and Stanley, J.C. (1966). Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally.

Casswell, Sally. (1982). The effect of pretesting on evaluation of a drug education program. Journal of Drug Education, 12, 173-80.

Cialdini, R.B., Petty, R.E., and Caciappo, J.T. (1981). Attitude and attitude change. Annual Review of Psychology, 32, 357-404.

Cook, T., and Campbell, D. (1979). Quasi-Experimentation: Design and Analysis Issues for Field Settings. Boston: Houghton-Mifflin.

Cornacchia, H.J., Bentel, D.J., and Smith, D.E. (Eds.). (1973). Drugs in the Classroom: A Conceptual Model for School Programs. Saint Louis: The C.V. Mosberg Co.

DiCostanzo, J.L. and Eichelberger, R.T. (1980). Reporting ANCOVA in evaluation settings. Evaluation Review, 4, 419-50.

Dixon, W.J. (Ed.). (1981). BMDP Statistical Software. Berkeley, California: University of California Press.

Drug Education is Linked to Use. (1972, December 3). New York Times, [citation incomplete]. In Berberian, R., Gross, G., Lovejoy, J., and Paparella, S. (1976). The effectiveness of drug education: A critical review. Health Education Monographs, 4, 377-98.

Drug Use Tied to Education: Study Shows Drug Abuse Increases with Education. (1972, December 1). Detroit Free Press [citation incomplete]. In Berberian, R., Gross, G., Lovejoy, J., and Paparella, S. (1976). The effectiveness of drug education programs: A critical review. Health Education Monographs, 4, 377-98.

Eagly, Alice and Himmelfarb, S. (1978). Attitudes and opinions. Annual Review of Psychology, 29, 517-54.

Ebel, H.C., Katz, D. and Rosen, A. (1975). Effects of a marijuana drug education program: A comparison of faculty-elicited and student-elicited data. Journal of Drug Education, 5, 77-85.

Einstein, S. and Allen, S. (1972). Proceedings of the First International Conference on Student Drug Surveys. Farmingdale, NY: Baywood Publishing.

Elashoff, Janet. (1969). Analysis of covariance: A delicate instrument. American Educational Research Journal, 6, 383-401.

Evans, R. and Raines, B. (1982). Control and prevention of smoking in adolescents: A psychosocial perspective. In Coates, T., Peterson, A., and Perry, C. (Eds.) Promoting Adolescent Health. New York: Academic Press.

Evans, S.H. and Anastasio, E.J. (1968). Misuse of ANCOVA when treatment effect and covariate are confounded. Psychological Bulletin, 69, 225-34.

Gasnier, Louis (Director). (1936). Reefer Madness. [film] Toronto: New Cinema Enterprises. [distributor]

Gilbert, J.P., Light, R.J. and Mosteller, F. (1975). Assessing social innovations: An empirical basis for policy. In Bennett, C.A. and Lumsdaine, A.A. (Eds.), Evaluation and Experiment, New York: Academic Press.

Goodstadt, M.S. (1975). The impact and roles of drug information in drug education. Journal of Drug Education, 5, 223-33.

Goodstadt, M.S., Cook, G. Magistad, S. and Gruson, V. (1978). The Drug Attitude Scale (DAS): Its development and evaluation. International Journal of the Addictions, 13, 1307-17.

Goodstadt, M.S., Sheppard, M.S. and Chan, G.C. (1982). Relationships betweeen drug education and drug use: Carts and horses. Journal of Drug Issues, 12, 431-442.

Gourlay, Neil. (1953). Covariance Analysis and Its Applications in Psychological Research. British Journal of Statistical Psychology, 6, 25-34.

Halpin, G. and Whiddon, T. (1977a). Drug education: Solution or problem? Psychological Reports, 40, 372-74.

Halpin, G. and Whiddon, T. (1977b). Relationships between drug attitudes and drug use. Measurement and Evaluation in Guidance, 10, 55-57.

Hanson, D.J. (1982). Alcohol and drug education. Journal on Alcohol and Drug Education, 2, 1-13.

Hochhauser, Mark (1979). Bias in drug abuse survey research. International Journal of the Addictions, 14, 675-87.

Hollander, M. and Davis, B. (1983). Trends in Adolescent Alcohol and Drug Use in Vancouver. Vancouver: B.C. Ministry of Health, Alcohol and Drug Programs.

Huck, S.W. amd McLean, R.A. (1975). Using Repeated Measures ANOVA to analyze the data from a pretest-posttest design: A potentially confusing task. Psychological Bulletin, 82, 511-518.

Kandel, Denise (1975). Stages in adolescent drug use. Science, 190, 912-914.

Kazdin, Alan. (1980). Research Design in Clinical Psychology. New York: Harper and Row.

Kinder, B.N., Pape, Nancy, and Walfish, S. (1980). Drug and alcohol education programs: A review of outcome studies. International Journal of the Addictions, 15, 1035-54.

Levy, M.R. (1971). Drug abuse education: A pedagogical schizophrenia. American Journal of Pharmacy, 143, 51-56.

Macro Systems, Inc. (1972). Evaluation of Drug Education Programs (Volumes 1-4). Unpublished manuscript. Cited in Goodstadt, M.S. (1975). Impact and roles of drug information in drug education. Journal of Drug Education, 5, 223-33.

Madsen, D.B. (1979). Group discussion and drug abuse prevention. International Journal of the Addictions, 14, 1117-23.

Meyers, J.L. (1979). Fundamentals of Experimental Design. Toronto: Allyn and Bacon.

National Commission on Marijuana and Drug Abuse. (1973). Drug Use in America: Problem in Perspective. Washington: U.S. Government Printing Office.

Nehemkis, A., Madari, M.A., and Lettieri, D.J. (Eds.). (1976). Drug Abuse Instrument Handbook. Rockville, Maryland: National Institute on Drug Abuse.

O'Rourke, T.W. (1980). Methodological considerations to improve anti-smoking research. Journal of Drug Education, 10, 159-69.

O'Rourke, T.W. and Barr, S.L. (1974). Assessment of the effectiveness of the New York State Drug Curriculum Guide with respect to drug attiudes. Journal of Drug Education, 4, 347-56.

Randall, Diana and Wong, M.R. (1976). Drug Education to Date: A Review. Journal of Drug Education, 6, 1-21.

Royse, D., Keller, S., and Schwartz, J.L. (1982). Lessons learned: The evaluation of a drug education program. Journal of Drug Education, 12, 181-90.

Schaps, E., Churgin, S., Palley, C., Takata, B., and Cohen, A.Y. (1980). Primary prevention research: A preliminary review of program outcome studies. International Journal of the Addictions, 15, 656-76.

Schaps, E., DiBartolo, R., Moscowitz, J., Palley, C. and Churgin, S. (1980). A Review of 127 Drug Abuse Prevention Program Evaluations. Unpublished manuscript. Lafayette, California: Pacific Insititute for Research and Evaluation.

Single, E., Kandel, D., and Johnson, B.D. (1975). The validity and reliability of drug use responses in a large scale longitudinal survey. Journal of Drug Issues, 5, 426-443.

Snider, J.G. and Osgood, C.E. (1969). Semantic Differential Technique. Chicago: Aldine Publishing Company.

Stuart, R.B. (1974). Teaching facts about drugs: Pushing or preventing? Journal of Educational Psychology, 66, 189-201.

Swisher, J.D., Warner, J.W. and Herr, E.L. (1972). Experimental comparisons of four approaches to drug abuse prevention among ninth and eleventh graders. Journal of Counselling Pychology, 19, 328-332.

Swisher, J.D., Crawford, J., Goldstein, R. and Yura, M. (1971). Drug education: Pushing or preventing? Peabody Journal of Education, 55, 69-75.

Thompson, Eva. (1978). Smoking Education Programs 1960-1976. American Journal of Public Health, 68, 250-57.

U.S. Department of Health Education and Welfare, Office on Smoking and Health. (1964). Surgeon General's Report on Smoking and Health. Washington: U.S. Government Printing Office.