Estimates of Intragroup Dependence for Drug Use and Skill Measures in School-Based Drug Abuse Prevention Trials: An Empirical Study of Three Independent Samples

Lawrence M. Scheier, PhD
Kenneth W. Griffin, PhD, MPH
Margaret M. Doyle, MPH
Gilbert J. Botvin, PhD

Group-randomized drug abuse prevention trials customarily designate schools as the unit of assignment to experimental condition; however, students within schools remain the unit of observation. Students nested within schools may show some resemblance based on common (peer) selection or school climate factors (i.e., disciplinary practices, group norms, or rules). Appropriate analyses of any treatment effects must be statistically correct for the magnitude of clustering within these intact social units (i.e., intraclass correlation coefficient [ICC]). There is little reported evidence, however, of variation in ICCs that might occur with studies of racially or geographically diverse populations. The purpose of this study was to generate estimates of intragroup dependence for drug use and psychosocial measures (hypothesized mediators) from three separate drug abuse prevention trials. Clustering for the drug use measures averaged .02 across study and age-groups (range = .002 to .053) and was equivalently small for the psychosocial measures (averaging .03 across studies and age-groups; range = .001 to .149). With few exceptions and across different samples, clustering decreased in magnitude over time. Clustering was largest for peer smoking and drinking norms among white, suburban youth and smallest for alcohol expectancies among urban black youth. Findings are discussed with respect to the influence of social climate factors and group norms in the design and analysis of school-based, drug abuse, prevention programs.

After almost a decade of declining prevalence rates, adolescent drug use has increased substantially since 1992.[1] Epidemiological reports based on nationally representative samples of secondary school students combined with evidence from national household surveys indicate marked upswings in alcohol, cigarette, and marijuana use. The increase in drug use has affected adolescents from different age and racial groups and has been observed for a variety of drugs. Initiation to drug use is occurring at increasingly younger ages and more youths are reporting problems associated with their drinking and drug use.

Lawrence M. Scheier, associate professor of counseling, University of Nevada, Las Vegas. Kenneth W. Griffin, assistant professor of psychology, Margaret M. Doyle, data analyst, and Gilbert J. Botvin, professor and director, Institute for Prevention Research, Division of Prevention and Health Behavior, Department of Public Health, Weill Medical College of Cornell University, New York.

Address reprint requests to Lawrence M. Scheier, Associate Professor, University of Nevada, Las Vegas, Counseling Department, 4505 Maryland Parkway, Campus Box 453045, Las Vegas, NV 89154; phone: (702) 895-3398; fax: (702) 895-1869; e-mail: lawrence.scheier@ccmail.nevada.edu.

Health Education & Behavior, Vol. 29 (1): 85-103 (February 2002)

The specter of increased morbidity and mortality related to early and continued drug use has helped to shape a public health agenda aimed at reducing teenage drug abuse.[2,3]

School-based drug abuse prevention programs represent a concerted response to the problem of adolescent drug abuse. Recent evidence points to the success of several theory-driven, school-based prevention programs in reducing cigarette, alcohol, and marijuana use.[4-11] A primary focus to determine the efficacy of these programs rests with measurement of change at the individual level.[12-13] Students are assessed periodically with respect to self-reported drug use as well as various skills and drug-related cognitions that are targeted causal agents of change. Despite the emphasis on assessing individual-level change, rarely are individual students the unit of assignment to treatment condition. Typically, in large-scale drug abuse prevention trials, whole schools are assigned to a treatment condition. When a sufficiently large number of schools are used, randomization at the level of the school affords a researcher the greatest level of control against potential biases and protects against threats to internal validity from potential confounds.[14-15]

Although randomization of schools to treatment conditions provides a simple remedy to a very complex problem, there are several design concerns associated with nested or multilevel data structures. First, when students are both the unit of assignment and unit of observation, there are usually considerable degrees of freedom (i.e., large sample sizes) and relatively few problems with obtaining sufficient statistical power for the comparisons of interest. However, when schools are the unit of assignment and the student is the unit of observation (i.e., program-related change is based on assessment of individual students' behavior), there is a considerable loss of statistical power because usually there are fewer schools than individuals for conducting statistical comparisons (and only very large effects will be noted). Researchers conducting prevention trials may have access to a large number of students but a relatively small number of schools that are available for random assignment to the experimental conditions.

A second concern is that students within schools tend to cluster together with respect to the outcomes of interest, and there exists higher resemblance of behaviors within as opposed to between schools. Possible mechanisms to account for the high levels of behavioral similarity include social contagion, peer selection, and peer socialization processes.[16-17] According to Ennett and colleagues, grouping in social networks is based on peer similarity and reinforcement through social interactions. Tightly knit groups of friends or social networks form based on similar activities and shared interests, which then provide a basis for enduring adolescent friendships. These friendships are created within natural boundaries formed by schools and facilitate social learning mechanisms that foster a wide range of behaviors. Empirical studies of peer social influence mechanisms document the importance of clusters or social cliques in the generation of antisocial behaviors.[18-20] Research also shows that an important feature supporting the strength of social groups rests with passive social influence processes that "involve social modeling of substance use by one's peers and friends, and one's perception of that use."[21] In this regard, intact social groups operatively defined as peer clusters help reinforce drug use behaviors by establishing normative beliefs that drug use is socially acceptable.

Prevention researchers are often faced with important design considerations in planning and evaluating school-based prevention trials. For example, failure to correct for the magnitude of clustering (i.e., intragroup dependence) by examining only individual-level data results in misspecification of the model and poses a threat to internal validity.[15] The most serious threat is a violation of the independence-of-errors assumption that is critical for conducting inferential statistical tests to determine treatment effects. Discounting the variation due to the group (i.e., school) biases standard errors of the treatment effect estimates and increases the effective Type I error rate (i.e., increases the probability that a researcher will reject a true null hypothesis[22]). On the other hand, restricting the analyses to include only group-level effects discounts important individual-level variation that may reflect hypothesized processes of individual-level change.[10] One effective solution to this problem is to compute the degree of relatedness of students within intact social groups (i.e., schools) and adjust any further statistical comparisons conducted at the individual level by the magnitude of intragroup dependence. The intraclass correlation coefficient (ICC) (designated as ρ) measures the magnitude of the variation in the data that is attributed to the unit of assignment (e.g., school) and can range from –1.00 to +1.00. The formula for computing ρ appears in several publications and is defined as the ratio of school-level variance to the total variance for the individual.[15,23] With posttest data from the typical school-based drug prevention study, the total variance in a particular outcome measure is a linear combination of two random components, school and residual error. The former component reflects variation due to the unit of assignment, whereas the latter component reflects residual variation due the unit of observation (i.e., individual-level variability). Depending on the nature of the research design, there also may be additional components of residual variation that may influence outcomes and can be controlled statistically (i.e., with multilevel or hierarchical data that are more complexly nested, additional variance components may be observed).

In addition to providing an estimate of the degree of relatedness among intact social groups, the ICC provides an important piece of information used to obtain correct sample size estimates. According to Donner, Birkett, and Buck,[24] computation of correct sample size estimates with multilevel or nested designs requires a modification to the variance of the estimated school means by a term called the *variance inflation factor* (VIF). VIF is a quantitative term reflecting the amount of underestimated variability in the outcome measure that can be attributed to clustering. In other words, VIF is the amount of variance adjustment made to the person-level variance to account for the magnitude of clustering within the data. VIF is calculated as follows: $\mathrm{VIF} = (1 + [n - 1]\rho)$, where n = the average number of students within schools (and may be computed as the harmonic mean when school sizes greatly vary) and ρ is the intraclass correlation. An increase in variance inflation can be expected when ρ positively deviates from zero and with increasing n. Thus, even trivial clustering (ICC) estimates with substantially large samples can reduce power. Computation of the ICC based on previous data collections conducted by the researcher or external estimates provided in the literature can help to estimate the loss of power, and researchers can profit from this information by including sufficient schools to restore power. These important design considerations have led several researchers to caution that "accurate sample size estimation thus requires the availability of precise ICC estimates for the outcome variables under study" (p. 426).[23]

Preliminary Evidence of the Magnitude of Intraclass Correlations

Despite the importance of considering intragroup dependence as a major design feature of school-based prevention trials, only a handful of studies have provided estimates of the magnitude of clustering effects for drug use outcomes.[19,22,23,25-28] Murray and Hannan[22] reported clustering estimates ranging from –.0002 to 0.048 obtained from a predominantly (91%) white sample. Interestingly, clustering estimates decreased in magnitude with the increasing prevalence of drug use observed between the 7th and 12th grade. Murray and Hannan suggest that high ICCs may reflect greater social cohesion when prevalence is low. In other words, peer selection mechanisms drive a select group of highly deviant youths to bond or cluster together. Despite the relatively small numbers of youths participating in this social grouping, their high-risk status accounts for a relatively large proportion of the total drug experience at that particular school. As drug use becomes more prevalent, the effect of clustering among a particular high-risk group diminishes, and no single identifiable clique can account for a widely distributed pattern of behavior. In addition, Murray and Hannan also reported substantial decreases in the magnitude of clustering estimates when homogeneous subgroups were created on the basis of urban versus suburban locale.

Murray and Short[28] provided clustering estimates for 9th- and 12th-grade students across 15 communities randomized to a community-based intervention or comparison group (i.e., individuals were nested within communities, which were nested within treatments). The intervention consisted of community-wide efforts to modify the physical, social, and policy environments of the participating communities in an effort to reduce underage drinking and drinking-related problems. Among 9th graders, clustering estimates ranged from a low of .002 for purchase attempts of alcohol in the past 30 days to a high of .019 for self-reported alcohol use in the past 30 days. Among 12th graders, clustering estimates ranged from a low of .003 for binge drinking in the past 2 weeks to a high of .022 for successful purchase attempts of alcohol in the past month. Interestingly, Murray and Short also report that clustering estimates showed considerable shrinkage in certain cases as well as substantial increases in other cases with application of regression adjustments. These covariate adjustments included controls for individual- (e.g., gender, age, and family living structure) and community-level influences (e.g., number of outlets selling alcohol, presence of a college in the surrounding area, and average unemployment).

Siddiqui and colleagues[23] provided clustering estimates for cigarette use and related outcomes obtained from an adolescent smoking prevention trial that randomized at the school level (program implementation was conducted at the classroom level). A cohort of multiracial students was surveyed at four time points from the seventh through ninth grades. At the classroom level, clustering estimates for current smoking status ranged from .04 (ninth grade) to .09 (seventh grade). Siddiqui et al. also noted that clustering effects were larger in magnitude for the behavioral and knowledge outcomes (i.e., health effects) compared to measures of coping, refusal skills, normative expectations, and behavioral intentions. Additional evidence also was provided that clustering effects were largest in magnitude for female students and largest overall for white students, followed by Hispanic and Asian youths and then black students.

In a separate study, Murray and colleagues[26] provided clustering estimates that were obtained from 11 smoking-prevention studies. Overall, clustering estimates ranged from an average of .006 (SD = .02) for a measure of the incidence of weekly smoking to .02 (SD = .01) for a measure of the prevalence of weekly smoking. Finally, in their analysis of social networks conducted with fifth- and sixth-grade elementary school students, Ennett and colleagues [19] reported clustering estimates for drug use measures ranging from a low of .012 for current marijuana use to a high of .051 for lifetime cigarette use.

Importance of the Current Study

A vital piece of information that can be gathered from additional empirical studies pertains to the generalizability of clustering estimates to different populations, different geographic regions, and different drug types. Much of the evidence gathered so far has been based on smoking prevention trials, and further information is needed to establish the magnitude of clustering with respect to alcohol and marijuana. In addition, there also is wide variability in prevalence rates based on race and geographic location. Urban ethnic minority youths, particularly black youths, report lower rates of drug use compared with white youths. Ethnic minority youths are disproportionately represented in metropolitan cities, and epidemiological evidence documents that drug prevalence rates considerably vary across rural, suburban, and inner-city schools.[1] Despite lower reported prevalence rates, minority youths report more alcohol-related problems in their neighborhoods and report observing more drug sales in their neighborhood than white students.[29] Environmental and contextual factors may play a large role in determining rates of drug use within a specific geographic area as well as within a specific school.[30-32] Gathering information on schools that represent different geographic locations should help provide a clearer picture of the influence of social climate and environmental features on rates of drug use.

A primary goal of this article then is to provide estimates of intragroup dependence from populations that are ethnically diverse and geographically unique. Two of the studies were conducted in minority-rich school districts in an effort to test the generalizability of a cognitive-behavioral, competence enhancement, drug abuse prevention program. All three studies contain moderately large numbers of schools and relatively large numbers of youths, all of which help to provide a basis for obtaining robust sample estimates. One study included schools from urban, suburban, and rural locations, and the remaining two studies gathered data from students attending inner-city schools. All three prevention trials included multiple follow-up assessments, which provide a basis for examining changes in clustering that may correspond to age (i.e., maturation) or other measurable factors. In addition to computing clustering estimates for the behavioral outcomes, clustering estimates also were examined for several targeted psychosocial risk mechanisms. To date, only Siddiqui et al.[23] and Murray and Short[28] have produced clustering estimates for psychosocial variables other than drug use outcomes.

METHOD

Randomization for each of the three school-based drug abuse prevention trials occurred at the level of the school; however, individual students within schools were the unit of observation (i.e., skills were taught to the individual students within classrooms and periodic assessments were made on the individual students). All three studies used the Life Skills Training Program as their focal prevention strategy.[33-35] A common set of research and implementation protocols guided the three prevention trials. These included (1) passive consent procedures, (2) unique identification codes to preserve confidentiality and facilitate longitudinal tracking, (3) procedures to enhance the validity of self-report data (carbon monoxide breath samples were collected from students), (4) questionnaires were completed during a single classroom period, (5) assessments were conducted annually from the seventh through ninth grade, (6) regular classroom teachers provided the intervention, (7) and project staff members from the research institute were used to monitor program fidelity and conduct the survey administration. Survey content for each of the three studies included behavioral items to tap frequency and intensity of alcohol, cigarettes, and marijuana use, as well as a wide range of cognitive, attitudinal, skills, and psychological variables hypothesized to foster initiation to drug use.

Study 1 consisted of a drug abuse prevention trial designed to follow a single cohort of primarily white, suburban, and rural middle school youths. A total of 56 schools from three geographic regions in the northeast portion of the United States participated in the prevention trial. The sample for Study 1 is 89.5% white and 52.4% male (the largest remaining ethnic group was 2.8% black). Seventy percent of the students reported living with both parents. The remaining students were broken down into 8.6% living with their mother and a stepparent, 15.3% living with their mother only, 1.8% with their father only, 2% with their father and a stepparent, and 1.9% living in some other situation. A pretest assessment was conducted in the fall of the seventh grade, and follow-up assessments were conducted annually thereafter through the ninth grade. Treated students received 15 intervention sessions in the seventh grade, 10 sessions in the eighth grade, and 5 booster sessions in the ninth grade. Further details on the conceptual scope of the intervention, selection of schools, assessment procedures, and randomization protocols can be found elsewhere.[4-5]

Study 2 was a drug abuse prevention trial conducted with inner-city Hispanic and black youths. A total of 47 schools (11 public and 36 parochial) from four boroughs in New York City participated in the study. The sample was 56% Hispanic, 19% black, 14% white, and 12% belonged to other race groups (Oriental or Asian). Slightly more than one-half (51%) of the students were female, and slightly more than one-half (58%) reported living with both their parents. A randomized block design was used, and schools were assigned to condition after blocking on school type and percentage of Hispanic students in each school. A total of 25 schools were assigned randomly to receive a 15-session Life Skills Training (LST) intervention, and 22 schools were designated as controls. Several publications detail the sample characteristics, study protocols, and intervention focus.[36]

Study 3 was a drug abuse prevention trial implemented with inner-city, ethnic minority youths to test the generalizability of findings obtained from Study 1 (a sample composed primarily of white, suburban youths). Prior to the beginning of the study in 1993-94, New York City schools with predominantly minority populations were assigned randomly to treatment ($N = 16$) and control conditions ($N = 13$) after blocking on a composite index of schoolwide cigarette smoking. Schools were carefully chosen based on district-wide information provided by the Board of Education regarding schools' Chapter 1 eligibility for government lunch subsidy and demographic information that detailed minority composition on a district basis. Invitations to participate in the prevention trial were sent to all superintendents, and shortly thereafter field staff met with individual principals within each district for recruitment purposes. All seventh graders attending schools assigned to the treatment groups took part in the LST drug abuse prevention program as a part of their regular curriculum. All treatment schools received additional prevention programming (booster sessions) in the eighth grade. Students attending schools assigned to the control condition received an information-only intervention as part of their regular curriculum. Control schools received additional information sessions in the eighth grade. At the pretest assessment, the sample was 48% male, 59% black, 24% Hispanic, and the remaining students were divided between Asian (6%) and non-Hispanic white categories (3%). Forty percent of the students reported living with both parents, 13% reported living with a parent and a stepparent, 36% reported living with their mother only, and 2% reported living with their father only. The remaining 9% reported some other living arrangement that did not involve a parent living in the home.

Behavioral and Psychosocial Measures

A total of 12 measures that were substantively linked to the intervention goals were selected for analysis in the present study. Drug outcome measures included frequency of alcohol use ("How often [if ever] do you drink alcoholic beverages?"), with response categories ranging from 1 (*never tried them*) through 9 (*more than once a day*); cigarettes ("How much do you generally smoke now?"), with response categories ranging from 1 (*never*) through 7 (*more than a pack a day*);* and marijuana use ("How often [if ever] do you usually smoke marijuana?"), with response categories ranging from 1 (*never tried it*) through 9 (*more than once a day*).

Hypothesized mediators represent several key psychosocial domains including drug-related cognitions (peer norms, expectancies, and knowledge), resistance skills (refusal efficacy and assertiveness), and personality (risk taking). In Studies 1 and 3, a three-item measure of refusal efficacy consisted of a single item to tap frequency of refusal assertiveness for smoking ("How often do you say 'no' when someone tries to get you to smoke?"), with response categories ranging from 1 (*never*) through 5 (*almost always*); a single item to tap frequency of refusal assertiveness for drinking ("How often do you say 'no' when someone tries to get you to drink?"), with an identical response format; and a single item to tap confidence in refusal skills ("Indicate how confident you are that you could do well in . . . refusing a cigarette offered by a friend"), with response categories ranging from 1 (*not at all confident*) through 5 (*very confident*). Internal consistency for this three-item scale based on Cronbach's alpha method was .73 in both studies. Study 2 used a slight variation of this scale and included a five-item scale assessing refusal coping methods. A common stem was provided ("If someone asked you to smoke, drink, or use marijuana or other drugs:"), and sample responses included "tell them not now" or "make up an excuse and leave." Response categories ranged from 1 (*definitely would*) through 5 (*definitely would not*). Internal consistency for the five items was .84.

Ten items to tap defense of rights, initiation, and assertiveness skills were taken from the Gambrill and Richey Assertiveness Inventory.[37] Sample items to tap defense of rights include "take something back to the store, if it doesn't work right," initiation includes "compliment your friends," and assertiveness skills include "express an opinion even though others may disagree with you." Response categories ranged from 1 (*never*) through 5 (*almost always*). Internal consistency for this scale across the three samples ranged from .58 to .82. Seven items were taken from the Eysenck Personality Inventory[38] to assess sensation seeking. Sample items include "I often wish I had more excitement" and "I get bored more easily than most people." Response categories ranged from 1 (*strongly disagree*) through 5 (*strongly agree*). Internal consistency for this scale based on the three samples ranged from .68 to .76.

Peer normative expectations were assessed with single items for drinking ("In your opinion, how many people your age drink alcoholic beverages?") and cigarette smoking ("In your opinion, how many people your age smoke cigarettes?"). Response categories for both items ranged from 1 (*none*) through 6 (*almost all*). Near and short-term health effects and factual knowledge regarding drug prevalence were assessed by a 10-item scale for drinking and separately a 10-item scale for cigarette smoking. Knowledge items

* The cigarette item for Study 2 varied slightly in wording from Studies 1 and 3 ("If you smoke cigarettes, about how much do you usually smoke?") and used an 8-point response format ranging from 1 (*none at all*) through 8 (*more than two packs a day*).

for alcohol included “A pregnant woman’s drinking can affect the health of her baby,” and a sample prevalence item was “Most adults drink alcohol everyday.” Knowledge of cigarette prevalence included “fewer people smoke now than 5 years ago,” and a sample health effect item was “Cigarette smoke contains a poisonous gas called carbon monoxide.” Response categories were scored dichotomously as “true” or “false.” A summary score reflecting the sum total of correct responses was created for the smoking items and separately for the drinking knowledge items. Study 2 focused largely on smoking prevention, and therefore, an alcohol knowledge scale was not included as a separate measure.

Separate 10-item scales were constructed to assess the perceived positive and negative social consequences (i.e., social reinforcement expectancies) for cigarette smoking and alcohol use. Sample alcohol expectancies included “If kids drink alcohol, it proves they’re tough” and “Kids who drink have more friends.” Sample cigarette expectancies included “Smoking cigarettes makes you look cool” and “Kids who smoke have more friends.” Response categories for both scales ranged from 1 (*strongly disagree*) through 5 (*strongly agree*). Internal consistency estimates across all three studies for the Smoking Expectancy Scale ranged from .73 to .79. Reliability for the Drinking Expectancy Scale for was .76 for both Study 1 and Study 3. Unfortunately, the focus on smoking prevention in Study 2 precluded having measures of alcohol expectancies and alcohol knowledge.

Analysis Strategy

For each of the three studies, estimates of the ICC were computed separately for the behavioral outcomes and hypothesized mediators for three time points, including the seventh-grade pretest (baseline) and two annual follow-up assessments. The Statistical Analysis System (SAS) (8.0) Proc Mixed procedure was used to derive variance estimates for the appropriate random (school) and residual components, and these variance estimates were then used to compute the ICCs.[39] In contrast with ordinary least squares (OLS) estimation, the SAS Mixed procedure uses a restricted maximum likelihood estimation (REML) method to derive variance components. REML produces more efficient fixed-effect estimates with unbalanced data and, more important, allows estimation of multiple components of variance (i.e., school and residual error term), whereas OLS permits estimation of only a single component of variance (i.e., residual error). Regression models for the eighth- and ninth-grade follow-up assessments were covariate adjusted to control for treatment effects (a binary measure designating treatment [1] and control [0]). Panel data were used for the two follow-up assessments (e.g., students providing data in the eighth grade also were present in the seventh grade, and likewise, students with ninth-grade data were present in the seventh and eighth grades).

RESULTS

Prevalence Rates for the Three Prevention Trials

For the purpose of establishing prevalence estimates, continuous frequency of drug use items were dichotomized into *ever use* and *nonuse* measures. Alcohol was the most prevalent drug used across all three prevention trials (20.5%, 14.2%, and 15.9% for the three studies, respectively, in the seventh grade; 43.3%, 22.5%, and 24.4% in the eighth grade; and 52.9%, 29.3%, and 28.4% in the ninth grade). Rates of reported alcohol use were noticeably higher among the predominantly white sample in Study 1 (20.5%, 43.3%, and 55.2%). Across all three studies, rates of drug use increased from the seventh to ninth grades, with the most dramatic increases occurring between the seventh and eighth grades. Cigarette use was highest in the predominantly Hispanic sample (Study 2). Across all 3 years and for each study, rates of reported cigarette use also increased dramatically (Study 1: 9.2%, 14.97%, and 19.3% for the seventh through ninth grades, respectively; Study 2: 18.4%, 22.5%, and 24.2%; and Study 3: 9.9%, 16.0%, and 16.1%). Rates of marijuana use were notably lower than prevalence estimates for alcohol but in some cases parallel those for cigarette use. There also were dramatic age-related increases in the percentage of youths reporting some experimental marijuana use (Study 1: 7.8%, 17%, and 24.5% for the seventh through ninth grades, respectively; Study 2: 0.9%, 3.5%, and 3.3%; and Study 3: 3.4%, 9.8%, and 12.5%, respectively). Reported level of marijuana use was notably higher among students in Study 1 (white) and relatively lowest among students in Study 2 (mostly Hispanic).

Although not presented in detail, prevalence estimates also were computed based on Race × Gender breakdowns for Study 2 and Study 3 (Study 1 was 90% white). Comparisons of these prevalence estimates indicate no discernible patterns (a complete set of tabled values can be obtained from the first author). Male and female students reported smoking cigarettes in equal proportions (27% and 22% in Study 2 for female and male students, respectively, and 21% and 18% in Study 3). The overall proportions of male and female students reporting drinking (47% and 49% for female and male students, respectively, and 31% for both male and female students in Study 3) and smoking marijuana (4.8% and 5.7% in Study 2 for female and male students, respectively, and 4% and 6.7% in Study 3) were similar. These patterns did not change appreciably over time, although the overall proportion of youths reporting some experimental drug use increased for both male and female students. Likewise, comparison of prevalence estimates by race groups (black versus Hispanic) did not show any appreciable differences. Cigarette use was reported by 23% and 27% of seventh-grade black and Hispanic youths, respectively, in Study 2 and 17% and 24% for black and Hispanic youths in Study 3. Alcohol was reported by 46% and 54% for black and Hispanic youths, respectively, in Study 2, whereas 37% and 30% of black and Hispanic youths reported alcohol use in Study 3. Marijuana use was reported by 6% for both black and Hispanic youths in Study 2 and 5.6% and 6.4% in Study 3 for black and Hispanic youths, respectively.

Table 1 provides ICC estimates for each of the three studies at each assessment point. Clustering estimates for the eighth and ninth grade are not covariate adjusted (except for treatment effects).* Overall, clustering effects for the behavioral measures were rela-

* We are grateful to the anonymous reviewers who suggested using model-free estimates rather than using covariate-adjusted models to generate clustering estimates. A number of studies have shown that adjustment for pretreatment measures reduces the magnitude of clustering by reducing the school component of variance in relation to the residual component.[15,22] Notwithstanding, model-free estimates can be used to estimate school sample sizes with the addition of a multiplier to adjust for covariances between pretreatment scores and posttest scores. In the traditional ANCOVA case, the term $1 - R^2$ is used to simulate covariances between pretreatment behaviors and posttest scores and can be extended to include a wide array of predictors including demographic characteristics (e.g., gender or ethnicity) that might predict posttest behaviors (see also Murray and Hannan[22] for a detailed listing of multiplier terms for the different types of analysis used in prevention studies). This term accurately reflects reductions to the total variance; however, in the case of a trial consisting of multiple random effects, variance reduction attributed to pretreatment behaviors is rarely shared evenly by the variance components (i.e., school and individual). In these cases, the reader is referred to Murray,[15] who provides computational methods to derive the adjusted member (e.g., individual) and group (e.g., school) variance terms.

(note continues on p. 95)

Table 1. Estimates of Intraclass Correlations (ICC) for Outcome and Mediator Variables: Results From Three School-Based Prevention Trials

	Study 1					
	7th Grade		8th Grade		9th Grade	
	$\hat{\rho}$	CI[a]	$\hat{\rho}$	CI	$\hat{\rho}$	CI
ALCFREQ	.033	.022-.055	.052	.034-.082	.043	.027-.070
CIGFREQ	.035	.023-.057	.018	.009-.032	.017	.008-.032
MARFREQ	.038	.025-.061	.022	.013-.039	.042	.026-.068
RISK	.021	.010-.039	.024	.012-.044	.007	–.001-.020
ASSERT	.010	.001-.026	.022	.010-.042	.012	.002-.028
REFUSE	.029	.018-.050	.025	.014-.043	.032	.019-.054
PEERSM	.149	.107-.215	.147	.105-.213	.092	.062-.141
PEERDR	.110	.073-.169	.108	.070-.169	.074	.043-.125
CIGEXP	.040	.025-.066	.052	.032-.084	.014	.004-.031
ALCEXP	.015	.002-.037	.037	.016-.072	.048	.023-.090
CIGKNOW	.006	–.004-.024	.034	.014-.068	.025[b]	.004-.054
ALCKNOW	.017	.008-.032	.014	.005-.029	.004	–.003-.016

	Study 2					
	7th Grade		8th Grade		9th Grade	
	$\hat{\rho}$	CI[a]	$\hat{\rho}$	CI	$\hat{\rho}$	CI
ALCFREQ	.026	.014-.047	.013	.004-.029	.039	.020-.074
GENSMOK	.008	.001-.021	.003	–.003-.014	.019	.005-.043
OFTPOT	.002	–.003-.011	.012	.003-.029	.018	.004-.042
RISK	.021	.009-.040	.006	–.001-.021	.022	.007-.048
ASSERT	.011	.002-.026	.022	.009-.045	.011	–.000-.033
REFUSE	.022	.010-.044	.014	.003-.034	.011	–.001-.033
PEERSM	.030	.017-.054	.040	.022-.070	.021	.007-.047
PEERDR	.026	.014-.048	.020	.008-.040	.022	.008-.049
CIGEXP	.014	.004-.030	.037	.019-.067	.017	.003-.042
ALCEXP	—[c]	—	—	—	—	—
CIGKNOW	.034	.019-.061	.101	.066-.159	.037	.017-.072
ALCKNOW	—[c]	—	—	—	—	—

	Study 3					
	7th Grade		8th Grade		9th Grade	
	$\hat{\rho}$	CI[a]	$\hat{\rho}$	CI	$\hat{\rho}$	CI
ALCFREQ	.011	.005-.025	.020	.010-.041	.024	.012-.051
GENSMOK	.035	.020-.066	.007	.002-.019	.013	.005-.031
OFTPOT	.053	.032-.097	.016	.008-.035	.013	.004-.030
RISK	.041	.017-.089	.018	.007-.042	.001	–.004-.012
ASSERT	.003	–.004-.012	.013	.006-.031	.014	.005-.034
REFUSE	.003	–.001-.012	.003	–.001-.012	.004	–.001-.015
PEERSM	.031	.018-.060	.012	.005-.027	.008	.001-.022
PEERDR	.020	.011-.041	.022	.011-.044	.003	–.001-.014
CIGEXP	.040	.024-.076	.000	–.003-.006	.001	–.003-.011
ALCEXP	.050	.030-.092	–.001	–.003-.005	.003	–.002-.014
CIGKNOW	.013	.006-.028	.024	.012-.048	.020	.009-.044
ALCKNOW	.015	.007-.032	.019	.010-.040	.017	.007-.039

Table 1 Continued

NOTE: Labels are the following: ALCFREQ = alcohol frequency; CIGFREQ = cigarette frequency; MARFREQ = marijuana frequency; RISK = risk taking; ASSERT = assertive skills; REFUSE = refusal efficacy; PEERSM = peer norms for cigarette use; PEERDR = peer norms for alcohol use; CIGEXP = cigarette social reinforcement expectancies; ALCEXP = alcohol social reinforcement expectancies; CIGKNOW = knowledge of health effects from cigarette use; ALCKNOW = knowledge of health effects from alcohol use. Samples sizes for the eighth- and ninth-grade data are based on merged cases present at the prior wave(s). Sample sizes: Study 1 (7th grade, N = 5,908; 8th grade, N = 4,809; 9th grade, N = 4,174); Study 2 (7th grade, N = 3,518; 8th grade, N = 2,763; 9th grade, N = 1,918); Study 3 (7th grade, N = 5,222; 8th grade, N = 4,181; 9th grade, N = 3,406). Harmonic mean number of students in each school in the 7th grade for Study 1 is 79, for Study 2 is 35, and for Study 3 is 94.

a. Confidence intervals were computed using the formula provided in Snedecor and Cochran.[42]

b. A negative (or near-zero) estimate was obtained for the school variance component for Time 4 cigarette smoking knowledge. Computations proceeded using the no-bound option in the Statistical Analysis System (SAS) PROC MIXED procedure to allow for negative variance estimates. In addition, the MIVQUE(0) (minimum-variance quadratic unbiased) estimation procedure replaced the restricted maximum likelihood estimation (REML) method. Murray[15] details how MIVQUE(0), a noniterative estimation procedure, provides robust estimates and is appropriate for large data sets as long as the residual component (individual variability) accounts for at least 90% of the total random variation.

c. Measure is not available in this study.

tively larger in Study 1 within each drug type across grades. Averaging across all three studies and within drug type, clustering effects were largest overall for alcohol (average across studies = .029 compared to .017 for cigarette and .024 for marijuana use). In Study 1, clustering estimates were relatively larger for marijuana compared to cigarette use, despite a smaller proportion of youths in this sample reporting marijuana use in the seventh and eighth grades. This pattern also was observed for students in Study 3 (inner-city black youths) and indicates that a small group of drug-abusing youths are responsible for establishing a normative climate and account for a large proportion of variation in reported levels of drug use.

In Study 2 (Hispanic sample), clustering estimates were largest for alcohol across all three grades, and in Study 3, clustering estimates were largest for marijuana in the seventh grade and then alcohol in the eighth and ninth grades. When average clustering estimates were computed separately for the behavioral and psychosocial measures and separately for the three studies, an interesting pattern emerged. Clustering effects for the psychosocial measures among the white suburban and Hispanic youths increased slightly between the seventh and eighth grades and then decreased in the ninth grade (.044, .051, and .034 for Study 1 and .023, .034, and .020 for Study 2, for all three grades, respectively). However, this same pattern was not observed for Study 3 (inner-city black youths), where clustering effects decreased steadily over time (.024, .012, and .008, respectively).

It is worth noting that the present study focuses on providing normative guidelines for clustering estimates encompassing a wide range of behaviors and skills common to school-based drug abuse prevention trials. Thus, it is perhaps more meaningful to generate unadjusted clustering estimates in the event that other researchers do not have repeated measures data or cannot adjust for similar covariates in their models. Overall, adjustments for prior behavior (i.e., drug use) decreased the magnitude of clustering estimates by no more than 1% to 3% for most of the measures across all three studies.

Clustering estimates within the study tended to be larger in magnitude for the psychosocial items (e.g., average clustering estimate for drug use was .033 for Study 1, whereas the clustering estimate for the psychosocial measures across grades was .043). In the case of students in Study 1, the larger magnitude for the psychosocial measures indicates that students reported greater similarity within schools on the basis of psychosocial functioning and less similarity with respect to drug use. This pattern was not, however, consistent across studies. In Study 2, composed primarily of Hispanic youths, average clustering estimate for the drug use measures was .016 and for the psychosocial measures was .026. However, in Study 3 (urban black youths), the average clustering estimate for drug use was .021, and the clustering estimate for the psychosocial measures was .015. Thus, there is less similarity in psychosocial functioning within schools and greater similarity or clustering in drug use among black youths. The smaller proportion of drug-abusing youths in Study 3, coupled with the relatively larger magnitude of clustering for drug use in this sample, would indicate the salient influence of a small collection of high-risk drug-abusing youths.

An examination of the patterns observed with regard to measures of psychosocial functioning bear out some of the evidence supporting the influence of small intact social groups. Within the different measures of psychosocial functioning, the largest degree of similarity was observed for those measures conceptualized as drug-related cognitions including peer smoking norms (Study 1), smoking knowledge (Study 2), and alcohol expectancies (Study 3). Average level of clustering for drug-related cognitions was .06 in Study 1 versus .02 for the skill measures, .03 for cognitions, and .01 for skills in Study 2, and .02 for cognitions and .01 for skills in Study 3.

Although not presented in their entirety, additional analyses were conducted to establish empirically whether clustering estimates differed based on gender, race, and geographic location. Tabled values for these comparisons are available from the first author. Briefly, these comparisons showed that clustering estimates differed based on gender both within and across studies. These differences were primarily observed for cigarette and marijuana use. Among the psychosocial mediators, clustering differed based on gender for peer smoking norms in Study 1. Across all three studies, clustering estimates for peer alcohol norms were larger in magnitude for female students compared with male students. Clustering estimates for risk taking in Study 2 (primarily Hispanic) were notably larger in magnitude for male students compared with female students, but this pattern did not hold in Study 3 (composed primarily of black youths).

Additional comparisons based on racial self-identification were conducted for Studies 2 and 3 (which contained sufficient representation of ethnic minorities to conduct valid comparisons). Although no consistent pattern emerged across the numerous comparisons, there was evidence that clustering effects for the drug use measures were notably larger for Hispanic youths with the exception of marijuana use, which was larger for black youths in both studies. Among the psychosocial measures, clustering effects also were larger for risk taking (i.e., sensation seeking) among Hispanic youths in both Studies 2 and 3.

Finally, Murray and Hannan[22] reported that geographic location (i.e., urban versus suburban) of schools made for large differences in the magnitude of clustering, with higher clustering estimates reported when schools are mixed in the same analysis. The present study tested the generalizability of those findings by distinguishing schools participating in Study 1 on the basis of urban ($N_{schools}$ = 4 and 3% of the sample, $\overline{X}_{harmonic}$ = 32.4), rural ($N_{schools}$ = 16 and 25% of the sample, $\overline{X}_{harmonic}$ = 69.4), and suburban

($N_{schools}$ = 36 and 72% of the sample, $\overline{X}_{harmonic}$ = 95.4) geographic location (where $\overline{X}_{harmonic}$ indicates the number of students per school). Although the results of these analyses are not presented in tabular form (and are available from the first author), there were several noted differences in the magnitude of clustering estimates based on geographic location. For example, clustering estimates for seventh-grade peer cigarette smoking norms were .217 for urban youths, .168 for suburban youths, and .047 for rural youths (by comparison, the magnitude of clustering based on the complete seventh-grade sample was .149). The disparity in magnitude for cigarette norms may, in part, be attributed to the relatively small number of urban youths comprising the overall sample. In effect, a relatively small and behaviorally homogeneous group of youths increases the magnitude of clustering. Accordingly, clustering estimates for the behavioral measures was relatively larger for urban youths (average for drug use measures was .09, .04, and .02 for urban, suburban, and rural youths, respectively). The average magnitude of clustering for psychosocial measures also was relatively larger for urban youths (.08, .05, and .03 for urban, suburban, and rural youths).

Clustering estimates for the drug use measures also were relatively larger for urban youths. Most notably, estimates for alcohol (.110) and marijuana use (.126) were considerably larger than either suburban (.032 and .046, for alcohol and marijuana, respectively) or rural youths (.028 and .010, respectively). Only one other discernible pattern was evident based on geographic location. With the exception of peer alcohol use norms, the magnitude of clustering for drug-related cognitions (norms, expectancies, knowledge) was relatively larger for urban youths, and there was little difference in these estimates for suburban and rural youths. As an added piece of information, proportional tests of independence indicated that drinking was more prevalent among suburban schools, $\chi^2(2)$ = 18.02, $p \leq .001$ (67% of suburban youths reported alcohol use vs. 29% and 3% of rural and urban youths, respectively), and, likewise, so was cigarette use, $\chi^2(2)$ = 24.73, $p \leq .001$ (63% of suburban youths reported cigarette use vs. 32% and 5% of rural and urban youths, respectively), and marijuana use, $\chi^2(2)$ = 18.13, $p \leq .001$ (64% of suburban youths reported marijuana use vs. 30% and 5% for rural and urban students, respectively).

DISCUSSION

Researchers conducting group-randomized drug abuse prevention trials can assign whole schools to treatment conditions in an effort to maximize control, limit bias, and protect internal validity. Despite assignment of schools to experimental condition, the determination of program efficacy still largely rests with analysis of individual-level outcomes. When schools are the unit of randomization and individuals are the unit of observation, the data are considered nested or hierarchical. An important feature of nested data is the observation that oftentimes students within the same schools are more homogeneous with respect to behaviors and skills than students in different schools. Schools represent an important source of social activity for students, and as a result, they may serve as vehicles for transmitting certain cultural mores that promote acquisition of behaviors and skills. The effect of students clustering within schools is to eliminate the independence between units of observation. In the event that a researcher applies inferential statistics to determine treatment group differences, failure to account for the presence of intragroup dependence (when ICC > 0) can produce biased (underestimated) standard errors for the fixed effects of interest and inflate the Type I error rate.[40]

Along these lines, several investigators have underscored the importance of providing good estimates of the magnitude of clustering to facilitate the design and planning of group-randomized trials.[23,26,28] To address this concern, we provide estimates of the magnitude of clustering obtained from three independent, school-based drug abuse prevention trials. The duration of each prevention trial, the use of a common core assessment strategy, the large numbers of schools, and the wide range of drug use and psychosocial measures help fashion this study as an important contribution to a recent, important, and rapidly growing literature. Without accurate estimates of clustering, prevention researchers must resort to simulation studies or guess to calculate power and sample sizes. The current study helps to alleviate this problem by providing valid estimates and augments previous reported empirical findings that were limited to smoking prevention studies.[23,26]

The prevalence of drug use across the three prevention trials was quite similar to national estimates obtained with similar age groups. The close symmetry between the rates provided in the present study and national estimates helps to rule out the possibility of historical influences on drug use that may have occurred across the different time frames these prevention trials were conducted. White and Hispanic youths consistently reported the highest rates of drug use for all three types of drug use. The relatively lower rates of drug use reported by black students may contribute to high levels of clustering, especially because the effect of small intact groups of high-risk youths within schools can appreciably inflate the magnitude of clustering.

More refined analyses based on racial breakdowns and using data from Study 2 and Study 3 indicated there were parallel trends in the proportion of youths reporting drug use over time. That is, when there was evidence of increasing numbers of youths reporting some drug use within a study, these increases also were observed for each race group across studies. However, Hispanic youths did somewhat outpace black youths in the use of drugs and by the ninth grade showed a disparity in their reported drug use compared with black youths. In effect, compared with their black counterparts, a greater number of Hispanic students reported using alcohol, cigarettes, and marijuana in Study 3 by the ninth grade. This trend was not evident in Study 2, despite a greater proportion of Hispanic youths purposively sampled.

For the most part, the size of the clustering effects was relatively small. However, most of the clustering estimates were significantly different from zero and indicate the importance of accounting for clustering effects in the analysis of group-randomized prevention trials. There were a few discernible patterns in the magnitude of clustering across the three prevention trials. Overall, clustering effects declined with age and were moderately larger for the hypothesized psychosocial mediators compared to the drug measures. Different factors may contribute to the decline in the magnitude of clustering over time. First, with increasing age, there were increases in drug prevalence rates, which may mean that drug use was becoming more normative and socially acceptable. In this respect, the influence of a few highly deviant and drug-using students would lessen on the larger social climate given that larger numbers of students reported experimental drug use. In other words, the effect of small clusters of students who smoke cigarettes, drink alcohol, or use marijuana at earlier ages diminishes considerably when the group norm shifts toward tolerating drug use. The same argument may hold true for the hypothesized mediators. It is important to recognize that many of the skills in question (e.g., assertiveness and refusal efficacy) are undergoing rapid transformation during this age period. As more and more youths acquire, test, and refine these skills, the effect of a few students with more, or conversely less, advanced skills should diminish. When these findings are viewed from a programmatic standpoint, intervention strategies that focus on improving social competence (i.e., refusal efficacy) can lead to improved skills among a wider audience of students. Likewise, prevention modalities that target correcting misperceptions of peer and adult drug use (i.e., normative beliefs) can promote more conservative estimates of drug use and thus weaken the power of small, albeit influential, social clusters.

Estimates of clustering for peer cigarette smoking and drinking norms were considerably higher among students in Study 1 than Studies 2 and 3. Study 1 was conducted with a primarily white, suburban sample, whereas Studies 2 and 3 were conducted with urban, ethnic minority youths. Differences of this nature have not been reported elsewhere in the literature and may attest to the varying strengths of peer groups in white versus minority communities. With respect to the remaining comparisons across the three prevention trials, there was little evidence of differential clustering based on race (when the prevention trials were compared at an aggregate level). The observation that the magnitude of clustering effects is minimally different across race groups has important ramifications for prevention. In particular, one issue raised by these findings is that despite implied cultural differences, the effect of small clusters of high-risk drug-abusing students as they contribute to establishing prevailing norms and social climates regarding drug use may be equally potent among inner-city ethnic minority youths as they are among suburban, white youths.

There also were some notable differences when comparisons were made across the three studies. For instance, in the transition between the seventh and eighth grades, clustering estimates increased for Study 1 and Study 2, and this was followed by a decrease in the magnitude of clustering estimates in the ninth grade. It may be that during the initial transition to middle school, a small intact clique of students who use drugs carries greater influence. As more and more youths experiment with drugs (as evidenced by the increasing prevalence rates across grades), there is some diminution of influence for these high-risk youths. However, this same pattern did not occur in Study 3, where clustering estimates decreased in magnitude steadily during the 3-year period.

Clustering estimates for skills and personality measures (measures hypothesized to mediate program effects) were relatively smaller than clustering estimates for drug-related cognitions (norms, expectancies, and knowledge measures). Norms (perceived beliefs regarding social acceptability of drug use and prevalence of drug use), knowledge (short- and near-term health effects), and expectancies (perceived positive benefits and negative consequences of drug use) reflect a generalized body of information obtained through vicarious learning and direct modeling experiences. A natural conduit for these experiences is formed through social clusters (i.e., close peer groups), which convey important social messages influencing the early stages of drug use. In contrast to cognitions, skills can take longer to accrue and may not be subject to influence from small pockets of students (i.e., those with accelerated levels of skills or those absent use of these skills).

Clustering analyses conducted but not reported in the present study also discerned important race and gender differences. For instance, clustering effects for risk taking among Hispanic youths were initially comparable or smaller than their black counterparts. However, with increasing age, the magnitude of these effects increased and so did the race disparity. Because clustering effects reflect the behavior of small intact social groups, these quantitative estimates may reflect the activity of a select few youths who are moving in the direction of unconventional behavior (i.e., sensation seeking). Moreover, the eighth grade seemed to reflect a focal point attesting to the strength of many of the psychosocial measures. That is, subgroup analyses based on race indicated that the magnitude of various clustering estimates increased in the eighth grade and paralleled rapid increases in rates of drug use. This pattern was notable for expectancies (.004 to .113 for black youths and –.002 to .038 for Hispanic youths in Study 2), cigarette smoking, and alcohol norms. This triumvirate of behavior, norms, and expectancies may reflect the consolidation of reinforcement contingencies among a select group of highly deviant youths. One final point worth noting with respect to observed race differences, clustering estimates for Study 3 in the ninth grade were noticeably small and in many cases for the psychosocial measures did not depart significantly from zero.

Gender-related analyses indicated greater coherence among female students in their behavior and psychosocial functioning. This was particularly true for cigarette use, where ICCs were larger for female compared with male students. Likewise, the magnitude of clustering for marijuana use in Study 3 (primarily black students) was larger for female students compared with male students (again attesting to the overall strength of a small cluster of black female students accounting for a large share of the behavioral uniqueness). Among the psychosocial measures, clustering was greater for risk taking among male students in Study 2 and greater among female students in Study 3. Estimates of clustering for peer smoking and alcohol norms also were considerably larger for female students compared with male students, and this pattern held up for each of the three prevention studies. Norms are a key ingredient in the etiology of early-stage drug use and, as indicated by the differential magnitude of clustering estimates for female and male students, may place certain girls at greater risk for later cigarette and other drug use.[41]

Using information from Study 1, which was conducted across diverse geographic locations, we also found notable differences between urban, suburban, and rural youths. Previously, Murray and Hannan[22] reported that clustering estimates dropped in magnitude when students were divided on the basis of urban versus suburban school location. If clustering reflects the amount of similarity when students are treated as a single entity within a school, then dividing the students into their respective "social" groups on the basis of geographic location should lessen the impact of a specific homogeneous lifestyle on the larger social climate. For instance, rural youths may have less access to drugs compared with urban youths, and when clustering is examined for these groups as a single quantitative estimate, the influence of one particular group can inflate upwardly the ICC to reflect passive social modeling (i.e., normative beliefs). Once this influence is removed by stratification, greater precision is obtained and leads to a concomitant reduction in clustering influence.

In the present study, students in suburban schools reported higher rates of all three types of drug use and were followed in decreasing order by rural and then urban youths. The magnitude of clustering was greatest overall for urban youths, who comprised a relatively small subset of the overall student population (and may reflect powerful clusters of students). This finding is consistent with those reported by Murray and Hannan,[22] who noted that among other factors, decreases in the size of social clusters are paralleled by increases in the magnitude of clustering effects. Among the hypothesized mediators, there also was evidence that school location factored into the magnitude of the ICC. Most notably, the size of the ICC for refusal skills among urban schools was almost twice the size of the clustering estimates for suburban and rural school students. In fact, 7 of the 12 estimates were considerably larger for the urban school students, where smaller numbers of students in each school portend a greater impact associated with social aggregations.[22]

Finally, it is also worth noting that there were a few clustering estimates that were negative. As a correlation statistic, the ICC can range between –1.00 and +1.00, where a negative sign indicates the absence of similarity within social units and a positive sign indicates similarity within groups. Virtually all the negative clustering estimates were close to zero, and the interpretation of this number would suggest that the degree of similarity among students within schools is less in magnitude compared to students between schools. In many instances, a lack of similarity between students within schools compared with that among students between schools is a chance occurrence. For researchers wishing to develop guidelines for interpreting clustering estimates and gauging these against normative standards for calculating appropriate sample sizes, a negative and small clustering estimate should be equated with zero and interpreted accordingly.

Implications for Practice

There are many uses for the ICC in the development and evaluation of effective drug abuse prevention interventions. First, the design effect or VIF is dependent on the computation of the ICC. Once the VIF has been estimated, researchers possess a quantitative handle that details how much they may possibly underestimate (or overestimate) the true intervention effect. Not only is the VIF sensitive to the magnitude of the ICC, but it also is sensitive to the numbers of schools. Thus, prior to mounting a costly and time-consuming prevention trial, a researcher can simulate or vary the ICC according to published guidelines based on external data and then vary the number of schools to provide a quantitative framework for estimating the precision of the presumed intervention effect. In other words, prior to actually conducting a field trial, prevention scientists can determine the required sample size that will help to minimize any variance inflation due to clustering within intact social groups.

Second, more often than not, practitioners are required to demonstrate that a particular intervention works before funding agencies will appropriate large sums of money for the conductance of large-scale field trials. In this regard, collection of pilot data is essential and mostly entails mounting a smaller scale prevention trial with fewer numbers of schools. It is still essential, however, to demonstrate program effectiveness using appropriate statistical tests that include some estimate of clustering effects. The provision of guidelines outlining the magnitude of clustering under different geographic, racial, and gender groupings provides an empirical basis from which to adjust individual-level error terms. Using this approach, an interventionist can determine more precisely the efficacy of an intervention and generalize empirical findings to similar settings that may contain a similar degree of clustering.

This study represents one of the first to generate clustering estimates for a wide range of behavioral and psychosocial measures and across three demographically diverse and racially heterogeneous samples. The inclusion of three relatively large, geographically, and racially diverse samples should provide ample basis from which to generalize to similarly constituted groups. Likewise, this study can provide sufficient information for other prevention researchers to conduct appropriate power analyses and sample size estimations. Where a researcher may not be fortunate enough to have sufficiently large numbers of schools to conduct multicomponent and multilevel analyses, the information provided in this study can be used to determine how much variance attenuation is required to correct for intragroup dependence. Such post hoc corrections including the VIF represent a viable means of correcting test statistics that are generated based on individual-level analyses, and while they may circumvent the unit-of-analysis problem that exists with group-randomized designs, they can provide a useful heuristic from which to gauge prevention effects.

References

1. Johnston LD, O'Malley PM, Bachman JG: *National Survey Results on Drug Use from the Monitoring the Future Study, 1975-1995. Vol. 1. Secondary School Students*. National Institute on Drug Abuse. Washington, DC, Government Printing Office, 1996. (NIH Pub. No. 96-4139.)
2. Office of National Drug Control Policy: *The National Drug Control Strategy, 1997*. Washington, DC, Executive Office of the President, 1997. (Document NCJ163915.)
3. Department of Health and Human Services: *Healthy People 2000: National Health Promotion and Disease Prevention Objectives: Full Report*. Washington, DC, Superintendent of Documents, 1990. (DHHS Pub. No. PHS 91-50212.)
4. Botvin GJ, Baker E, Dusenbury L, Tortu S, Botvin EM: Preventing adolescent drug abuse through a multimodal cognitive-behavioral approach: Results of a 3-year study. *J Consult Clin Psychol* 58:437-446, 1990.
5. Botvin GJ, Baker E, Dusenbury L, Botvin E, Diaz T: Long-term follow-up of a randomized drug abuse prevention trial in a white middle-class population. *JAMA* 273:1106-1112, 1995.
6. Botvin GJ, Dusenbury L, Baker E, James-Ortiz S, Kerner J: A skills training approach to smoking prevention among Hispanic youth. *J Behav Med* 12:279-296, 1989.
7. Ellickson PL, Bell RM: Drug prevention in junior high: A multi-site, longitudinal test. *Science* 247:1299-1305, 1990.
8. Ellickson PL, Bell RM, Harrison ER: Changing adolescent propensities to use drugs: Results from Project ALERT. *Health Educ Q* 20: 227-242, 1993.
9. Hansen WB, Graham JW: Preventing alcohol, marijuana, and cigarette use among adolescents: Peer pressure resistance training versus establishing conservative norms. *Prev Med* 20: 414-430, 1991.
10. MacKinnon DP, Johnson CA, Pentz MA, Dwyer JH, et al: Mediating mechanisms in a school-based drug prevention program: First-year effects of the Midwestern Prevention Project. *Health Psychol* 10:164-172, 1991.
11. Pentz MA, Dwyer JH, MacKinnon DP, Flay BR, et al: A multicommunity trial for primary prevention of adolescent drug abuse: Effects on drug use prevalence. *JAMA* 261:3259-3267, 1989.
12. Donaldson SI, Sussman S, MacKinnon DP, Severson HH, Glynn T, Murray DM, Stone EJ: Drug abuse prevention programming: Do we know what content works? *Am Behav Scientist* 39:868-883, 1996.
13. MacKinnon DP, Weber MD, Pentz MA: How do school-based drug prevention programs work and for whom? *Drugs and Society* 3:125-143, 1989.
14. Issacs S, Michael WB: *Handbook in Research and Evaluation*. San Diego, CA, EdITS, 1990.
15. Murray DM: *Design and Analysis of Group-Randomized Trials*. New York, Oxford University Press, 1998.
16. Ennett ST, Bauman KE: Adolescent social networks: School, demographic, and longitudinal considerations. *J Adolesc Res* 11:194-215, 1996.
17. Kandel DB: Processes of peer influences in adolescence, in Silbereisen RK, Eyferth K, Rudinger G (eds.): *Development as Action in Context: Problem Behavior and Normal Youth Development*. New York, Springer-Verlag, 1986, pp. 203-228.
18. Aloise-Young PA, Graham JW, Hansen WB: Peer influence on smoking initiation during early adolescence: A comparison of group members and group outsiders. *J Appl Psychol* 79:281-287, 1994.
19. Ennett ST, Flewelling RL, Lindrooth RC, Norton EC: School and neighborhood characteristics associated with school rates of alcohol, cigarettes, and marijuana use. *J Health Soc Behav* 38:55-71, 1997.
20. Sussman S, Dent CW, Stacy AW, Burciaga C, Raynor A, et al: Peer-group association and adolescent tobacco use. *J Abnorm Psychol* 99:349-352, 1990.
21. Graham JW, Marks GS, Hansen WB: Social influence processes affecting adolescent substance use. *J Appl Psychol* 76:291-298, 1991.

22. Murray DM, Hannan, PJ: Planning for the appropriate analysis in school-based drub-use prevention studies. *J Consult Clin Psychol* 58:458-468, 1990.
23. Siddiqui O, Hedeker D, Flay BR, Hu FB: Intraclass correlation estimation in a school-baesd smoking prevention study: Outcome and mediating variables, by sex and ethnicity. *Am J Epidemiol* 144:425-433, 1996.
24. Donner A, Birkett N, Buck C: Randomization by cluster: Sample size requirements and analysis. *Am J Epidemiol* 114:283-286, 1981.
25. Hedeker D, Gibbons RD, Flay BR: Random-effects regression models for clustered data with an example from smoking prevention research. *J Consult Clin Psychol* 62:757-765, 1994.
26. Murray DM, Rooney BL, Hannan PJ, et al: Intraclass correlation among common measures of adolescent smoking: Estimates, correlates, and applications in smoking prevention studies. *Am J Epidemiol* 140:1038-1050, 1994.
27. Norton EC, Bieler GS, Ennett ST, Zarkin GA: Analysis of prevention program effectiveness with clustered data using generalized estimating equations. *J Consult Clin Psychol* 64:919-926, 1996.
28. Murray DM, Short B: Intraclass correlation among measures related to alcohol use by school aged adolescents: Estimates, correlates and applications in intervention studies. *J Drug Educ* 26:207-230, 1996.
29. National Institute on Drug Abuse: *National Household Survey on Drug Abuse: Race/Ethnicity, Socioeconomic Status and Drug Abuse: 1991*. Rockville, MD, U.S. Department of Health and Human Services, 1993.
30. Simcha-Fagan O, Schwartz, JE: Neighborhood and delinquency: An assessment of contextual effects. *Criminology* 24:667-703, 1986.
31. Skager R, Fisher DG: Substance use among high school students in relation to school characteristics. *Addict Behav* 14:129-138, 1989.
32. Wills TA, Pierce JP, Evans RI: Large-scale environmental risk factors for substance use. *Am Behav Scientist* 39:808-822, 1996.
33. Botvin GJ: Principles of prevention, in Coombs RH, Ziedonis, DM (eds.): *Handbook on Drug Abuse Prevention: A Comprehensive Strategy to Prevent the Abuse of Alcohol and Other Drugs*. Boston, Allyn & Bacon, 1995, 19-44.
34. Botvin GJ: Substance abuse prevention through life skills training, in Peters R, McMahon RJ (eds.): *Preventing Childhood Disorders, Substance Abuse, and Delinquency*, Banff International Behavioral Science Series, Thousand Oaks, CA, Sage, 1997, pp. 215-239.
35. Botvin GJ, Dusenbury L: Life skills training: A psychoeducational approach to substance abuse prevention, in Maher CA, Zins JE (eds.): *Psycoheducational interventions in schools: Methods and procedures for enhancing student competence*. New York, Pergamon, 1987, pp. 46-65.
36. Botvin GJ, Dusenbury L, Baker E, James-Ortiz S, Botvin EM, Kerner J: Smoking prevention among urban minority youth: Assessing effects on outcome and mediating variables. *Health Psychol* 11:290-299, 1992.
37. Gambrill ED, Richey CA: An assertion inventory for use in assessment and research. *Behav Ther* 6:550-561, 1975.
38. Eysenck HJ, Eysenck SBG: *Manual of the Eysenck Personality Questionnaire*. London, Hodder and Stoughton, 1975.
39. Littell RC, Milliken GA, Stroup WW, Wolfinger RD: *SAS System for Mixed Models*. Raleigh, NC, SAS Institute Inc., 1996.
40. Donner A, Koval JJ: Design considerations in the estimation of intraclass correlation. *Ann Hum Genet* 46:271-277, 1982.
41. Hansen WB, Graham JW, Sobel JL, Shelton DR, et al: The consistency of peer and parent influences on tobacco, alcohol, and marijuana use among young adolescents. *J Behav Med* 10:559-579, 1987.
42. Snedecor GW, Cochran WM: *Statistical Methods* (8th ed.). Ames, Iowa State University Press, 1989.