Meta-Analysis and Models of Substance Abuse Prevention

Betsy Jane Becker

INTRODUCTION

The idea of synthesizing available information about treatment efficacy or the strength of relationships among variables is not new. Procedures for combining such evidence date to the 1930s (Fisher 1932) and have been widely applied in the social sciences since Glass' introduction of meta-analysis in the 1970s (Glass 1976; Glass et al. 1981).

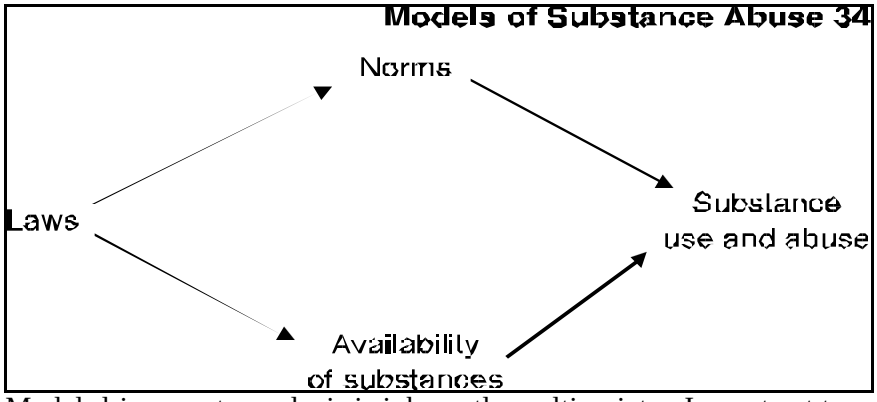
Recently reviewers in a number of disciplines have realized that research synthesis techniques can be applied in reviews of issues more complex than those previously studied. Meta-analysis has been criticized for attending only to main effects (Cook and Leviton 1980) and ignoring the important roles of mediating and moderating variables. Applications of meta-analytic techniques to complex processes (Becker 1992*b*; Premack and Hunter 1988), as well as methodological developments (Becker 1992*a*, 1992*c*), show that this oversimplification need not occur (see also Cook et al. 1992, p. 341).

This chapter introduces research synthesis methods for the analysis of complex processes and outlines how they can be applied in the study of the literature on substance abuse prevention. In particular, the chapter describes a model for the roles of risk and protective factors in substance abuse prevention, based on the review of Hawkins and colleagues (1992). The author next discusses how evidence about models could be gathered and examined in a quantitative synthesis of the literature on this topic, and describes key issues that arise in the application of this approach. A brief example of data analysis for a four-variable model is also presented. The chapter concludes with a discussion of how a model-based synthesis of risk and protective factors could be used in the design and analysis of substance abuse prevention programs.

MODEL-DRIVEN META-ANALYSIS

Model-driven meta-analysis refers to the quantitative synthesis of evidence pertaining to a model of the interrelationships among a set of constructs or variables. Often such models are illustrated using flowcharts or path diagrams. Flay and Petraitis (1991) showed two very detailed models of behavior that have served as theoretical frameworks for drug use behavior. Flay's model focused primarily on the psychological antecedents of drug use, whereas Elliott and colleagues (1985) outlined a broader sociological model for delinquent behavior.

Figure 1 shows a simple diagram of the roles of three broad social context factors influencing substance abuse (variables are drawn from Hawkins et al. 1992). Models can show direct influences, such as the relationships of norms and availability to substance use and abuse shown in figure 1. Indirect relationships (mediated by other variables) can also be shown. Laws are depicted in figure 1 as having two indirect influences on abuse.



Model-driven meta-analysis is inherently multivariate. In contrast to narrative reviews and more limited syntheses of bivariate relationships, model-based meta-analysis can provide quantitative evidence about interactive effects of relevant variables. This should be particularly useful in a review of evidence on drug abuse, since "[T]here is little evidence available regarding the relative importance and interactions of

various risk factors in the etiology of drug abuse" (Hawkins et al. 1992, p. 65). Similarly, Flay and Petraitis noted that, despite many reviews of correlates of drug use, "[T]here is no information about the relationships among the correlates" (1991, p. 82). Under certain assumptions, it may be possible to examine a model through meta-analysis that yields information about interactions not tested in any primary research study. Those assumptions are described more fully below.

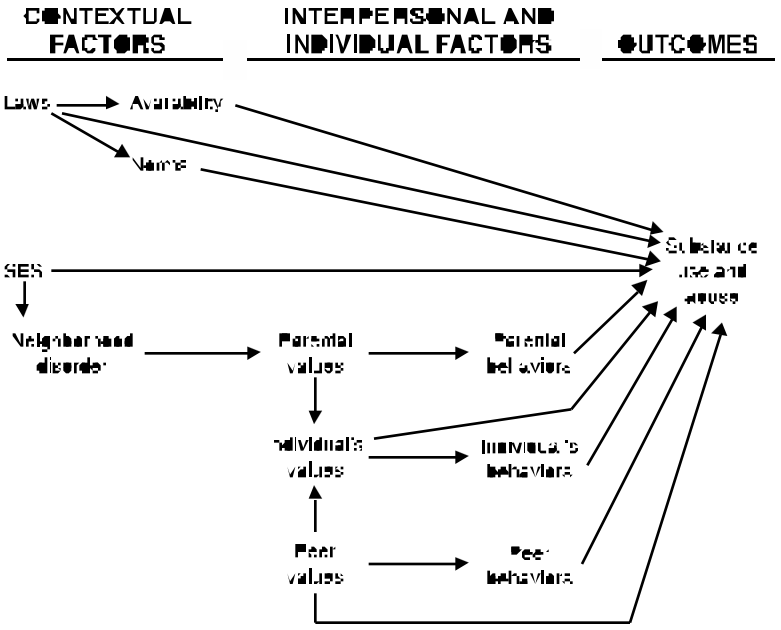
The models examined in a model-driven synthesis may arise empirically or be derived from theory. The theoretical-empirical distinction is rarely clear cut. Empirical research arises from implicit models of theory, and theory is often modified or even "discovered" by empirical work. An empirical model shows relationships that have been examined in primary research. This chapter describes an empirically derived model based on the narrative review of Hawkins and colleagues (1992). However, several authors (including Hawkins and colleagues) have noted the importance of a theoretical model or "conceptual framework for evaluating the content of substance abuse prevention curricula" (Hansen 1992, p. 408). Flay and Petraitis (1991) described 12 ways that theory is important in the area of substance abuse.

Theoretically derived models provide a context in which to assess the existence and the strength of evidence about a proposed model. Some parts of a theoretical model (e.g., hypothesized relationships) may be well studied, whereas others may never have been studied. These less studied (or unstudied) aspects of a model may be appropriate domains for further research. Clearly, it will be difficult or impossible to conduct a compre-hensive model-driven quantitative synthesis of a process if the bulk of the relationships proposed by the model have not been studied. Data requirements are discussed below.

Finally, model-driven meta-analyses can provide reviewers and policy-makers with information about processes that can help in practical decisions and program design. For instance, a review of the process of substance abuse may identify influences or combinations of influences that could be targeted in a substance abuse prevention program. Derivation of an empirical model may even allow the reviewer to test particular ideas about program features.

MODELS OF THE ROLES OF RISK FACTORS IN SUBSTANCE ABUSE

Models of the roles of risk and protective factors in substance abuse are implicit in the narrative review by Hawkins and colleagues (1992). Figure 2 shows one possible model that incorporates contextual factors and many of the individual and interpersonal factors described in the review.[1] The model shown in figure 2 has 11 broad predictors of substance use and abuse outcome for a total of 12 components. Table 1 lists those components.



Five components represent contextual factors, while the rest are interpersonal (parent and peer) and individual factors. The outcome itself is broadly defined, and leads to a good example of how such process models can be further delineated. For example, one could refine the model in figure 2 by focusing on drug abuse or on alcohol abuse. Some

TABLE 1. *Components in the model of substance abuse.*

Components	Examples
Laws Norms Availability Neighborhood disorder Socioeconomic status (SES)	
Peer values	Advocacy of drugs
Peer behavior	Drug and alcohol use, aggression, acceptance of individual
Parental values	Permissiveness towards drugs, educational aspirations for children
Parental behavior	Drug and alcohol use, hostility, marital dissolution, family conflict
Individual values	Attachment to parents, liking of school, educational expectations
Individual behavior	Delinquent behaviors, aggression, school performance, intellectual ability
Substance use and abuse	

predictive factors may be more relevant for one outcome than another; factors that are irrelevant to a particular outcome could be omitted from the refined model for that outcome.

The model in figure 2 shows 19 paths or connections between components. Both direct and indirect influences are outlined. Another way of refining the model is to change the paths shown in the model. For instance, all three "values" components have both direct and indirect connections to substance use and abuse. A different model might remove the three direct paths and show only indirect influences (i.e., those moderated by relevant variables). Moreover, this model does not show parent, peer, and individual behaviors. Such relationships may be important, but they are not direct paths to the outcome.

The model in figure 2 is certainly not the only possible model of the process described by Hawkins and colleagues. It is not an exact representation, since some of the component factors are very broad. However, it illustrates one process model that could be examined in a model-driven synthesis, and is based on empirical evidence.

EVIDENCE IN A MODEL-DRIVEN META-ANALYSIS

Existence of Research

The model in figure 2, having been derived from a narrative review of existing work, is empirical. Hawkins and colleagues cite evidence about many of the relationships shown in figure 2. Table 2 shows counts of the etiological studies reviewed by Hawkins and colleagues for each of the paths or components listed in table 1 and depicted in figure 2. Table 2 is an example of the first kind of evidence provided by a model-driven synthesis: existence of research on particular relationships.

These totals are based on each relationship described by Hawkins and colleagues (1992) and categorized according to the two components listed in table 1 that best matched the interrelated variables. Studies that examined several risk factors were included for each relationship studied. Original primary research was not consulted for this coding; decisions were made on the basis of the brief descriptions in Hawkins and colleagues' (1992) report. Different classification decisions could have been reached either with more information about the studies or by a coder more familiar with the literature on substance abuse.

Table 2 also includes studies that examined relationships on paths not depicted in figure 2; these counts are underlined. Five direct relationships not shown in figure 2 were examined in studies cited by Hawkins and colleagues (1992). Additionally, six entries represent relationships (denoted by asterisks) described as potentially important by Hawkins and colleagues and shown in figure 2, but not examined by any etiological studies in their table of results. Thus 6 of 19 paths, or nearly one-third of the paths in figure 2, are apparently unstudied. If this model is truly representative of the process of substance abuse development, research is needed to understand these paths in the model.

Several trends are apparent in table 2. First, the bulk of the studies mentioned in Hawkins and colleagues' (1992) table 1 looked at direct relationships of predictor variables to the substance use/abuse outcome. Of the 192 relationships counted, 149 (78 percent) involved substance use or abuse. Also, nearly half of the use/abuse relationships involved individual factors as predictors (i.e., the individual's own values and behaviors). Parental factors were mentioned next most frequently; 37 studies (roughly 25 percent) of use/abuse outcome examined parental values and behaviors as predictors. Finally, of the 43 instances in which the relationship did not involve the focal use/abuse outcome, over 90 percent (39 instances) were relationships in which the individual's behaviors (other than use/abuse) were the outcome.

Many of the possible entries in table 2 are simply empty. These empty positions represent paths that neither appear in figure 2 nor are mentioned by Hawkins and colleagues. As noted above, alternative models might include those other paths, and it is likely that studies not reviewed by Hawkins and colleagues (1992) included examination of those paths.

In an actual model-driven meta-analysis, thorough searches would be conducted to identify studies relevant to all paths in the model or models. Searches for model-driven meta-analyses often involve more extensive keyword lists and search strategies than more traditional meta-analyses or narrative reviews (Becker 1992*b*).

Analysis of Existing Data

Table 2 shows counts of studies that examined relationships relevant to the proposed model of substance abuse in figure 2. Many of the studies included in these counts probably presented their results in terms of indexes of association. In a quantitative synthesis of the evidence concerning the substance abuse model, the reviewer would retrieve and analyze these measures of association. Analyses of those measures provide the second type of evidence in a model-driven meta-analysis: evidence about strengths of relationships. These analyses are discussed in the following section.

ISSUES IN THE SYNTHESIS OF DATA

Cooper (1989) outlined five stages in the research synthesis process: problem formulation, data collection, data evaluation, data analysis, and reporting of results. Both problem formulation and data collection have been briefly discussed above. Problem formulation deals primarily with selecting or deriving a model or models to study. Data collection (gathering of studies) in a model-driven synthesis is

likely to be more extensive than that for a traditional quantitative review, as mentioned above, because of the multivariate nature of model-driven syntheses.

Data Evaluation

Cooper's third stage, data evaluation, involves retrieving study outcomes and coding study features such as study quality and characteristics of samples, measures, and, possibly, treatments. Coding study quality is at least as important in model-driven synthesis as it is in a more traditional meta-analysis. Also in a model-driven synthesis, the reviewer must code information relevant to the models being studied. For example, the studies reviewed by Hawkins and colleagues (1992) were classified according to the paths in the hypothetical model. This step would be crucial in a more extensive review because incorrect or careless classification could prompt critics to argue that dissimilar studies (apples and oranges) had been combined.

Between-Studies Differences. Because meta-analyses have often been criticized for overlooking important between-studies variables, coding these variables is critical. Differences in samples (e.g., age or SES of subjects), in the nature and duration of treatments given, and in study quality can all lead to variation in results.

Variation in outcomes (e.g., the strength of relationship of particular predictors with substance abuse) sometimes can be explained by a small number of between-studies variables (that is, study-level covariates). Then fixed-effects models may apply, and the relevant study features may be moderator variables for one or several paths in a model.

In other cases, between-studies variation may not be accounted for even after many study characteristics have been examined. In these cases, random-effects models may be applied. Essentially, the reviewer expects some uncertainty or amount of variation across studies or uncertainty in the strengths of relationships studied. One assumes that different populations (or more precisely, populations with different correlation structures) may have been examined in different primary studies. The object is to estimate variability or uncertainty in the population correlations and to incorporate those estimates into further analyses of the data. The distinction between fixed and random models has both conceptual and statistical subtleties (for more information on fixed- and random-effects models see Hedges 1994). Both fixed- and random-effects approaches are available for the synthesis of model-based data (Becker 1992*c*).

Data Retrieval. At the data evaluation stage, the reviewer retrieves correlational (associational) data from the primary research. While

correlation indices such as Pearson's r, Spearman's rho, and the phi coefficient are often reported, many studies yield more complex data. Regression analyses, canonical correlation analysis, and path analyses may provide data on relationships of interest, but their results are not as easily synthesized as zero-order correlations. The object of data retrieval is to retrieve the same index of association from each study (for each relationship) or to convert the indices that are retrieved into values that are comparable across studies. Often the correlation is the most useful index (i.e., most easily made comparable).

Specific illustrations are easily found in the literature on substance abuse. Extensive research by Brook and colleagues (1983, 1986) has examined the correlates of adolescent drug use. However, though zero-order correlations of many predictors to the drug use outcome are presented in some studies (e.g., Brook and colleagues 1986), intercorrelations among the predictors are not given directly but are incorporated into canonical correlation analyses. Another format for presentation of information on drug use correlates is found in Brook and colleagues (1983). Drug users were first categorized by level of drug use, then mean values for each of the correlates were reported for each group of users.

An additional data retrieval issue concerns the measurement of the substance use and abuse outcome. If a study measures substance use as a dichotomy, typical measures of association that assume bivariate normality of both variables (e.g., Pearson's r) are inappropriate. However, it may be possible to convert more appropriate measures for association (given this dichotomy) into indices of the correlation between continuous variables that might underly the dichotomy. McDermott (1984) examined the associations among parental drug use (measured as use versus nonuse), attitude toward adolescent drug use (categorized as permissiveness versus disapproval), and adolescent drug use (also measured dichotomously); three 2 x 2 tables presented the categorical results. In a more complex analysis of a dichotomous alcohol use outcome, Barnes and Welte (1986) used discriminant analysis to relate more than 10 potential predictors to alcohol use. Indices from studies with dichotomous outcomes will also differ in their statistical properties from those based on continuous outcomes, such as level or amount of substance use. At present, the methodology for synthesizing model-based results for dichotomous outcomes has not been developed.

When different studies report results of analyses of different statistical models (i.e., models that control for different factors), they provide information about different partial relationships. Thus the slope for peer drug use from a regression of SES, parental drug use, and peer drug use on child’s use of drugs is not comparable to the slope for peer drug use when SES is the only other predictor. Analyses of

structural models and regression models often pose this difficulty. Hansen and coauthors (1987) examined an elaborate hierarchical model of drug use using structural equation analyses. Their extensive results included reports of many path coefficients and residual correlations, but no zero-order correlations. Many other similar examples exist, suggesting that research on how to handle indices of partial relationships may be an area for further inquiry. Combining estimates of different parameters (e.g., of relations under different model specifications) is not sensible and is likely to yield inconsistent results in many circumstances. Combining zero-order indices avoids this confusion.

Data Analysis

For simplicity, temporarily assume that a reasonable number of studies have been gathered that examine all or parts of a proposed model. Further, assume that the studies provide zero-order correlation indices for the relationships studied. Several questions can then be posed about the relationships under study.

Procedures for analyzing correlational data in model-driven meta-analysis are described elsewhere (Becker 1992*c*; Becker and Schram 1994). The methods require that zero-order correlations be presented in each primary research study, or that they be retrievable from other study indices.

These methods enable the reviewer to ask, first, whether all studies show the same pattern of interrelationships among the variables in a correlation matrix (here, among the 12 components in the model). Then the reviewer can estimate a common correlation matrix (if studies appear similar) or a pooled matrix that accounts for between-studies variation in the correlation of values. Finally, either of these average matrices can be used to estimate standardized regression models showing the relative importance of the different predictors as well as intercorrelations among them. The reviewer can then piece together, from an entire literature, models similar to path-analytic (causal) models derived in single studies. The potential of these procedures to elucidate the nature of complex processes is tremendous. The approach has both strengths and weaknesses, however, as described below.

Availability of Data.

As described in the section on data evaluation, obtaining zero-order correlations or measures of association is necessary to apply currently available model-based synthesis methods. However, many studies do not present complete correlation matrices or indices of zero-order relationships.

Indices of partial relationships present problems of comparability, as discussed above. Missing (unreported) or nonexistent data cause statistical problems in estimation of average correlation matrices across studies (Becker 1992*a*). Data are considered missing, for instance, when a researcher reports correlations for a set of predictors with a substance use/abuse outcome but does not report intercorrelations among the predictors that also appear in the model under study.

If a proposed path in a model has not been studied, an average correlation for that path cannot be estimated. This may lead to a misspecified model if the omitted predictor is crucial. Estimated effects for studied variables may be biased if an important variable is omitted from the model. Such model misspecification can lead to incorrect conclusions, but may be difficult to avoid when using existing primary research. This problem highlights the importance of thoroughly searching for and collecting relevant studies.

Between-Studies Differences. Between-studies differences in study features as well as in the nature and extent of reported data may also present problems in a model-driven synthesis. Consider a very simple illustration by returning to the model in figure 1. Suppose that the search had identified 50 studies relevant to the four paths in figure 1, but that half of the studies examined adolescent drug use and half studied adults. Further suppose that these two groups of users are known to differ dramatically in many ways. If all of the studies of adolescents had examined the relationships of laws to norms and norms to use/abuse, and studies of adults had examined the remaining paths, it would not be possible to generalize about the entire model from the studies. Usually, the situation is not so clearly confounded as in this illustration.
If the reviewer is willing to apply a random-effects conceptualization to the model, however, some conclusions can be drawn. This is equivalent to arguing that, although the particular groups studied may represent different populations (e.g., of user types), there exists a "population of user populations" that is of interest. The task then is to determine how different the patterns of relationships appear to be in the populations being considered.

Artifactual Variation. Another source of between-studies differences in results that poses a problem in meta-analysis is artifactual variation. This can include such influences as differential reliability of measures (even if identical constructs have been studied) and restriction of range. For instance, results based on samples drawn from a single population can differ if one sample is unselected and the other is composed of high scorers (e.g., selected on the basis of an employment selection test or other similar instrument).

Corrections for both unreliability and range restriction are readily available for a series of single (bivariate) correlations (e.g., Schmidt and Hunter 1994). However, until recently the effects of applying these corrections to correlation matrices had not been studied. Schram (1995) examined a variety of methods for correcting correlation matrices for attenuation due to unreliability, and found that the familiar univariate correction performed well. Schram also derived a large-sample variance-covariance estimator for the corrected matrix that incorporates uncertainty due to the estimation of both the correlations and the reliability coefficients. While the reviewer may not have access to complete information about artifacts, it is important to acknowledge that artifactual variation can lead to variation in observed results.

Causality. When models are used in the planning of substance abuse programming, there is an implicit assumption that manipulation of relevant predictor values can lead to changes in substance abuse. Essentially, program planners are looking for potential causal relationships. Strong inferences of causality require both temporal precedence of the cause relative to the effect and elimination of other competing explanations of change in the outcome. Cook (1990, 1991) has written extensively about causality in meta-analysis and program evaluation.

An Example

To illustrate the possibilities for quantitative synthesis of correlational data, an example is presented of a synthesis of results from three samples. These three samples all arise from a single study by Mills and Noyes (1984). This is an overly simplistic example that avoids issues such as differential unreliability, comparability of constructs, and range restriction that might arise in a more realistic example.

The three samples are of 8th, 10th, and 12th graders from Maryland public schools. Four "use" variables are examined, two of which will be treated as predictors (smoking and use of alcohol) and two as outcomes (use of marijuana and cocaine). Methods used are described in Becker (1992*c*) and Becker and Schram (1994).

Table 3 shows the upper halves of the correlation matrices for the three grades. Each sample provides six correlation values. The first task is to ask whether the three sets of correlations arise from a single population. If so, a single pooled correlation matrix can adequately represent relationships in all the samples.

Example. The test of whether a single population correlation matrix applies to the three grades is a chi-square test with (3-1) x 6 = 12 degrees of freedom. For the data in table 3 the value is 25.64, which is significant at $p < 0.025$. The results to not appear to be completely consistent with the model of a single underlying population correlation structure. Thus a random-effects model can be adopted and an average correlation matrix can be estimated.

Estimating Variation in Population Correlations. In order to incorporate the uncertainty or variation in correlation strength that results from having samples from several populations with different correlation structures, the variances (and covariances) among the population correlations must first be estimated. Becker and Schram (1994) describe how the estimation and maximization (EM) algorithm can be applied to obtain these variance component estimates. For each relationship, an estimate is obtained of the variation in the population values of correlations representing that relationship.

For example, let $r_{SA(i)}$ represent the correlation between levels of smoking and alcohol use in study i. Then $\$_{SA(i)}$ is the corresponding population

TABLE 3. *Correlation matrices from Mills and Noyes (1984).*

	Smoking	Alcohol	Marijuana	Cocaine
Grade 8 (N = 672)				
Smoking	1.00	0.48	0.50	0.20
Alcohol		1.00	0.50	0.24
Marijuana			1.00	0.37
Cocaine				1.00
Grade 10 (N = 691)				
Smoking	1.00	0.43	0.54	0.25
Alcohol		1.00	0.52	0.27
Marijuana			1.00	0.43
Cocaine				1.00
Grade 12 (N = 589)				
Smoking	1.00	0.34	0.38	0.23
Alcohol		1.00	0.43	0.20
Marijuana			1.00	0.42
Cocaine				1.00

correlation. The model for a single correlation value under random effects shows that

$$r_{SA(i)} = \$_{SA(i)} + e_{SA(i)},$$

and thus variation (uncertainty) in the sample values of r_{SA} incorporates variation in the $\$_{SA(i)}$ values and sampling variance. The variance component for the smoking/alcohol use correlations is an estimate of variation among the $\$_{SA(i)}$ values. Similarly, covariances are estimated among the population correlations. If $\$_{SM(i)}$ represents the correlation of smoking with marijuana use in population i, the covariance component for $\$_{SA}$ and $\$_{SM}$ would be Cov ($\$_{SA(i)}$, $\$_{SM(i)}$) across populations.

Example. For the data from Mills and Noyes (1984), the EM algorithm produced a variance covariance matrix (denoted T) for the six correlation indices of

	$\$_{SA}$	$\$_{SM}$	$\$_{SC}$	$\$_{AM}$	$\$_{AC}$	$\$_{MC}$	
$\$_{SA}$	.0027	.0028	-.0003	.0016	.0010	-.0007	
$\$_{SM}$	.0028	.0038	.0001	.0022	.0015	-.0003	
$\$_{SC}$	-.0003	.0001	.0003	.0000	.0001	.0003	= T.
$\$_{AM}$	.0016	.0022	.0000	.0013	.0009	-.0002	
$\$_{AC}$	.0010	.0015	.0001	.0009	.0007	-.0000	
$\$_{MC}$	-.0007	-.0003	.0003	-.0002	-.0000	.0004	

Variances are shown on the diagonal, and the covariances are the off-diagonal elements of T. The standard deviations of the six populations of correlation values are 0.052, 0.062, 0.017, 0.036, 0.025, and 0.020. The second correlation, representing the relationship between smoking and marijuana use, shows the most variation. A standard deviation of 0.062 would correspond to a normal distribution ranging roughly from -0.20 to 0.20, if centered on zero. Even this is not a broad range for correlations.

Estimating the Average Correlation Matrix. Once an estimate of variation in the population correlations has been obtained, it can be incorporated in the estimation of an average correlation matrix. The estimate of the mean correlation matrix is obtained via generalized least squares (GLS) estimation (Becker 1992*c*). The GLS estimates can be obtained under fixed- and random-effects models. Covariation among the several correlations from each sample is accounted for in both cases. In the random-effects model, variation and covariation in population effects are also incorporated into the uncertainty of the estimates.

The random-effects GLS estimate of the mean correlation matrix for the three samples is

	Smoking	Alcohol	Marijuana	Cocaine
Smoking	1.00	0.42	0.48	0.23
Alcohol use	0.42	1.00	0.49	0.24
Marijuana use	0.48	0.49	1.00	0.41
Cocaine use	0.23	0.24	0.41	1.00

.

Comparing this estimate with the original data matrices in table 3 shows that the values of the sample correlations of smoking with alcohol use (r_{SA}) and with marijuana use (r_{SM}) indeed vary more about these means than the other correlation values, as suggested by the variance components in T above.

The variance-covariance matrix for the set of six average correlations is

	$r_{\_SA}$	$r_{\_SM}$	$r_{\_SC}$	$r_{\_AM}$	$r_{\_AC}$	$r_{\_MC}$
$r_{\_SA}$	0.0012	0.0010	-0.0000	0.0007	0.0004	-0.0002
$r_{\_SM}$	0.0010	0.0016	0.0002	0.0008	0.0006	-0.0001
$r_{\_SC}$	-0.0000	0.0002	0.0005	0.0001	0.0002	0.0003
$r_{\_AM}$	0.0007	0.0008	0.0001	0.0007	0.0004	-0.0000
$r_{\_AC}$	0.0004	0.0006	0.0002	0.0004	0.0007	0.0002
$r_{\_MC}$	-0.0002	-0.0001	0.0003	-0.0000	0.0002	0.0005

.

As could be expected from the amount of variation in the population values, the averages of the first two correlations, $r_{\_SA}$ and $r_{\_SM}$, show the most uncertainty, with standard errors of 0.035 and 0.040, respectively. These are still quite small relative to the magnitudes of the average correlations, however, which are both about 0.40.

Estimating Linear Models. Once an estimate of a mean correlation matrix has been obtained, it can be used to estimate a variety of predictive models for the intercorrelated variables. Here two standardized regression models are estimated. The first incorporates smoking and alcohol use as predictors of marijuana use.

The estimated model for this regression (based on the random-effects mean correlations and their variance, given above) is

$$\hat{M} = 0.33\ S + 0.35\ A,$$

where ^M represents a predicted standardized (z) score on the marijuana use scale, S is a z score for level of smoking, and A is a z score for level of alcohol use. The slopes, their standard errors, and tests of the hypothesis = 0 for each slope are given in table 4. Both slopes differ significantly from zero at very stringent levels. It is also possible to test whether the two slopes (say, $_S$ and $_A$) are equal, using their variances and the estimated covariance between b_S and b_A. The test of H_0: $_S = _A$ uses the statistic:

which has a standard normal distribution when H_0 is true. Since _ z _ œ 1.96, it is not significant at the = 0.05 level. It can be concluded that both level of smoking and level of alcohol use are significant, and equally strong, predictors of marijuana use.

The second model examines smoking, alcohol use, and marijuana use as predictors of cocaine use. The estimated model is

$$\hat{C} = 0.03\ S + 0.05\ A + 0.37\ M,$$

where ^C is a z score for level of cocaine use and S, A, and M are as described above. Table 4 shows that the only significant predictor in this

TABLE 4. *Standardized regressions showing contributions to substance abuse*.

	Outcome					
	Marijuana use			Cocaine use		
Predictor	b	SE(b)	z	b	SE(b)	z
Smoking	0.33	0.028	11.79*	0.03	0.039	0.82
Alcohol use	0.35	0.033	10.61*	0.05	0.049	0.92
Marijuana use	---	---	---	0.37	0.048	7.73*

NOTE: All predictors and outcomes represent standardized scores on the four "use" variables.

KEY: * = significant slope coefficients.

model is (standardized) level of marijuana use. Levels of smoking and alcohol use do not predict level of cocaine use for these 8th through 12th graders.

Display of Regression Results. Figure 3 shows the results of the standardized regression analyses displayed on a flow diagram similar to those in figures 1 and 2. The slopes are entered on the paths in the model, and significant slopes are starred. This model shows that across the three grades there are direct relationships between smoking and marijuana use, alcohol and marijuana use, and marijuana use and cocaine use. The effects of smoking and alcohol use on cocaine use are only indirect (i.e., mediated by level of marijuana use). According to the tests in table 4, the two paths (from S to C and from A to C), representing direct effects of smoking and alcohol use on cocaine use, could be eliminated.

Summary of Example. This example indicates the possibilities for analyses when results from multiple studies (here samples) are combined using techniques for model-driven quantitative synthesis. Tests of homogeneity (consistency) of results indicate whether fixed- or random-effects models are most appropriate. Average correlation matrices can be inspected for their own intrinsic value or used to obtain estimates of the

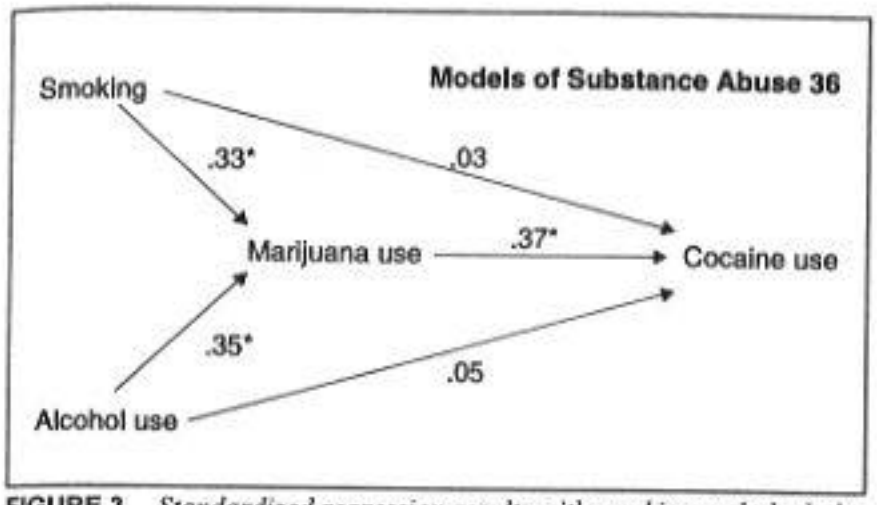


FIGURE 3. *Standardized regression results with smoking and alcohol use as predictors of marijuana and cocaine use.*

simultaneous relationships of several predictors to each outcome. These analyses can then inform the reviewer about the plausibility of a variety of models of relationships among variables.

USING MODELS FOR PROGRAM PLANNING

Most researchers and practitioners dealing with substance abuse prevention use both theory and empirical research in program planning. Reviews of school-based abuse curriculums (e.g., Hansen 1992) and other prevention programs (Tobler 1986, 1992) emphasize these ideas. Flay and Petraitis (1991) also discuss the importance of theory for program planning.

Hansen (1992) devoted nearly one-fifth of a review to the conceptual underpinnings of curriculum content for school-based programs, and described "the building block theoretical concepts used by researchers" and the "theoretical or quasi-theoretical assumptions about the means by which [program] components affect behavior" (1992, p. 408). Hansen's framework "provides a description of programmatic approach linked to mediating process" (1992, p. 408). A quantitative model-driven meta-analysis can provide an empirical assessment of proposed models such as those described by Hansen.

Tobler's two reviews (1986, 1992) also describe mediating processes underlying program strategies (or modalities). Table 1 in each article describes the assumptions of five program strategies. For instance, peer programs assume that "peer

pressure can impact attitudes and behaviors" (1992, p. 6). In the model shown in figure 2, these assumptions could refer to the peer values, individual values, individual behavior predictors, and the substance use/abuse outcome. The premises underlying Tobler's two types of peer programs involve different beliefs about the kinds of individual responses (behaviors) that can inhibit drug use. Model-based meta-analyses with sufficient data can support detailed comparisons of those program types, or of their assumptions. Such comparisons may aid in the refinement of existing program designs or the development of new programs that incorporate strategies that seem to work better in combination than in isolation.

Status Studies Versus Intervention Studies. One question the reviewer must address in conducting a model-based meta-analysis is whether to include both intervention studies and "status studies" in which no manipulation of variables is attempted. If both are included, it will be important to examine differences between the results of the two kinds of studies. The presence of an intervention could attenuate relationships seen in a one-group status study (of the same relationship) by making subjects appear to be more similar on the manipulated variable than they would naturally be. Alternately, if an intervention differentiates subjects (e.g., by making them more variable on coping skills, self-esteem, or knowledge of the effects of drug use), a study of that intervention may show a stronger relationship of the manipulated variable to substance use and abuse than a status study. Thus, intervention studies may not present the same view of the potential effectiveness of intervention strategies as status studies.

Comparisons of Program Models. To be most useful for program planning, a model-driven meta-analysis should contrast and compare different process models. Does a model that includes components for both peer and parental behaviors explain more variation in substance use and abuse than one dealing with peers only? Is attention to the individual's values necessary to understand levels of substance use/abuse? These questions imply different process models and different program designs.

Perhaps parental behaviors explain considerable variation in child drug abuse, but securing parental interest and participation in a substance abuse prevention program may be both difficult and costly. With a model-driven synthesis, such practical questions about program design can be weighed in light of concrete evidence about differences in process models.

Decisions about which models can or should be compared can be based on theory or on the need to make specific decisions about program components.[2]

CONCLUSION

This chapter illustrates the potential of model-driven quantitative synthesis for exploring and testing models of the influences on substance abuse, and for providing information for substance abuse prevention program planning. The application of these ideas in a thorough empirical review of the literature provides an exciting possibility for future work.

NOTES

1. Physiological factors have been omitted from this model. Other parts of the model are greatly simplified by creating very broad categories (e.g., "behaviors" and "values"). Other more differentiated models (e.g., specifying and separating particular behaviors) are possible.

2. Clearly, if the collection of studies for the meta-analysis does not include data on the models of interest, such comparisons will be impossible. The above discussion assumes that sufficient data are, in fact, available.

REFERENCES

Barnes, G.M., and Welte, J.W. Patterns and predictors of alcohol use among 7-12th grade students in New York State. *J Stud Alcohol* 47(1):53-62, 1986.

Becker, B.J. "Complications and Corrections in Constructing Synthetic Regressions." Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1992*a.*

Becker, B.J. Models of science achievement: Factors affecting male and female performance in school science. In: Cook, T.D.; Cooper, H.; Cordray, D.S.; Hartmann, H.; Hedges, L.V.; Light, R.J.; Louis, T.A.; and Mosteller, F., eds. *Meta-Analysis for Explanation: A Casebook.* New York: Russell Sage Foundation, 1992*b*.

Becker, B.J. Using results from replicated studies to estimate linear models. *J Educ Stat* 17:341-362, 1992*c*.

Becker, B.J., and Schram, C.M. Examining explanatory models through research synthesis. In: Cooper, H., and

Hedges, L.V., eds. *The Handbook of Research Synthesis.* New York: Russell Sage Foundation, 1994.

Brook, J.S.; Whiteman, M.; and Gordon, A.S. Stages of drug use in adolescence: Personality, peer, and family correlates. *Dev Psychol* 19:269-277, 1983.

Brook, J.S.; Whiteman, M.; Gordon, A.S.; and Cohen, A.S. Some models and mechanisms for explaining the impact of maternal and adolescent characteristics on adolescent stage of drug use. *Dev Psychol* 22(4):460-467, 1986.

Cook, T.D. Meta-analysis: Its potential for causal description and causal explanation within program evaluation. In: Albrecht, G., and Otto, H., eds. *Social Prevention and the Social Sciences: Theoretical Controversies, Research Problems, and Evaluation Strategies.* New York: Walter de Gruyter, 1991.

Cook, T.D. The generalization of causal connections: Multiple theories in search of clear practice. In: Sechrest, L.; Perrin, E.; and Bunker, J., eds. *Research Methodology: Strengthening Causal Interpretations of Nonexperimental Data.* Conference proceedings, U.S. Dept. of Health and Human Services, May 1990.

Cook, T.D., and Leviton, L.C. Reviewing the literature: A comparison of traditional methods with meta-analysis. *J Pers* 48:449-472, 1980.

Cook, T.D.; Cooper, H.; Cordray, D.S.; Hartmann, H.; Hedges, L.V.; Light, R.J.; Louis, T.A.; and Mosteller, F., eds. *Meta-Analysis for Explanation: A Casebook.* New York: Russell Sage Foundation, 1992.

Cooper, H.M. *Integrating Research: A Guide for Literature Reviews.* Beverly Hills, CA: Sage, 1989.

Elliott, D.S.; Huizinga, D.; and Ageton, S.S. *Explaining Delinquency and Drug Use.* Beverly Hills, CA: Sage Publications, 1985.

Fisher, R.A. *Statistical Methods for Research Workers.* 4th ed. London: Oliver and Boyd, 1932.

Flay, B.R., and Petraitis, J. Methodological issues in drug use prevention research: Theoretical foundations. In: Leukefeld, C.G., and Bukoski, W.J., eds. *Drug Abuse Prevention Intervention Research: Methodological Issues.* National Institute on Drug Abuse Research Monograph 107. DHHS Pub. No. (ADM)91-1761. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1991.

Glass, G.V. Primary, secondary, and meta-analysis of research. *Educ Res* 5(10):3-8, 1976.

Glass, G.V.; McGaw, B.; and Smith, M.L. *Meta-Analysis of Social Research.* Beverly Hills, CA: Sage, 1981.

Hansen, W.B. School-based substance abuse prevention: A review of the state of the art in curriculum, 1980-1990. *Health Educ Res* 7(3):403-430, 1992.

Hansen, W.B.; Graham, J.W.; Sobel, J.L.; Shelton, D.R.; Flay, B.R.; and Johnson, C.A. The consistency of peer and parent influences on tobacco, alcohol, and marijuana use among young adolescents. *J Behav Med* 10(6):559-579, 1987.

Hawkins, J.D.; Catalano, R.F.; and Miller, J.Y. Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychol Bull* 112:64-105, 1992.

Hedges, L.V. Statistical considerations in problem formulation. In: Cooper, H., and Hedges, L.V., eds. *The Handbook of Research Synthesis.* New York: Russell Sage Foundation, 1994.

McDermott, D. The relationship of parental drug use and parents' attitude concerning adolescent drug use to adolescent drug use. *Adolescence* 73:89-97, 1984.

Mills, C.J., and Noyes, H.L. Patterns and correlates of initial and subsequent drug use among adolescents. *J Consult Clin Psychol* 52(2):231-243, 1984.

Premack, S.L., and Hunter, J.E. Individual unionization decisions. *Psychol Bull* 103:223-234, 1988.

Schmidt, F.L., and Hunter, J.E. Correcting for sources of artifactual variation across studies. In: Cooper, H., and Hedges, L.V., eds. *The Handbook of Research Synthesis.* New York: Russell Sage Foundation, 1994.

Schram, C.M. "A Comparison of Methods for Correcting Multivariate Data for Attenuation, with Application to Synthesizing Correlation Matrices." Ph.D. diss., Michigan State University, 1995.

Tobler, N.S. Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcome results of program participants compared to a control or comparison group. *J Drug Issues* 16(4):537-567, 1986.

Tobler, N.S. Drug prevention programs can work: Research findings. *J Addict Dis* 11(3):1-28, 1992.

AUTHOR

Betsy Jane Becker, Ph.D.
Professor
Michigan State University
456 Erickson Hall
East Lansing, MI 48824-1034

Click here to go to page 120