December 1997
Version 6

Covariates of Alcohol Dependence and Abuse: A Multivariate Analysis of a 1988 General Population Survey in the United States

Bengt O. Muthén

Graduate School of Education & Information Studies

UCLA

Los Angeles

California 90024-1521

U.S.A.

This research was supported by grant AA 08651-01 from the National Institute on Alcohol Abuse and Alcoholism for the project "Psychometric Advances for Alcohol and Depression Studies." Jin-Wen Hsu, Ginger Goff, and Li-Chiao Huang provided helpful research assistance. I am grateful to Dr. Bridget Grant, Dr. Deborah Hasin, and Dr. Thomas Harford for their comments on an earlier version of this paper. I thank Linda Muthén for improving the clarity of the presentation.

Covariates

Abstract

This paper studies alcohol dependence and abuse criteria using data from a large general population survey in the U.S., the 1988 National Health Interview Survey (NHIS88). The NHIS88 collected responses to questions on alcohol consumption, family history of alcoholism, and a set of symptom items designed to capture DSM-III-R and the proposed DSM-IV criteria for diagnosing alcohol dependence and abuse. In previous work by Muthen, Grant & Hasin (1993), a two-dimensional model was found for these criteria. The two dimensions correspond to the level of severity of the criteria and may be seen as a new way to define abuse and dependence. In the present paper, the two-dimensional model is used to study the relationships between the criteria and covariates such as alcohol consumption, family history of alcoholism, and socio-demographic characteristics. External validity of alcohol abuse or dependence diagnoses in the general population is usually studied in a risk factor framework using logistic regression. In that case, a diagnostic variable is regressed on a set of covariates or hypothesized antecedents. This paper suggests an alternative to this approach that uses a structural equation modeling framework. This alternative is suggested because it may offer more generality in understanding the external validity of diagnoses. The structural equation analysis gives further support for the two dimensions underlying the criteria in that the two factors have different relationships to the background variables. In particular, the more severe factor is more strongly related to family history of alcoholism.

Key Words: Alcohol dependence, DSM-III-R, diagnostic criteria, risk factors, structural equation modeling.

Introduction

In response to continuing evaluations, both conceptual and empirical, the Diagnostic and Statistical Manual of Mental Disorders (DSM; American Psychiatric Association, 1987, 1992) for alcohol use disorders has undergone several revisions in the past decade. There is a need for research to understand the structural relationships between the criteria set forth to measure alcohol abuse and dependence. Muthén, Grant, and Hasin (1993) used factor analysis to study the dimensionality of the diagnostic criteria for alcohol abuse and dependence included in the 1988 National Health Interview Survey (NHIS88). The NHIS88 is a U.S. general population survey on alcohol use and alcohol-related problems which includes a set of symptom items developed to measure the DSM-III-R and the proposed DSM-IV diagnostic criteria for alcohol abuse and dependence. Muthén et al. (1993) found that a two-dimensional model fit the criteria well. In this paper, data from the NHIS88 will be used to study the external validity of this two-dimensional model.

The two-dimensional model of alcohol abuse and dependence represents a departure from unidimensional diagnostic schemes related to DSM-III-R and DSM-IV. Because of this, it is important to study the validity of the model using information external to the diagnostic criteria. This paper will relate the dimensions to covariates such as alcohol consumption, family history of alcoholism, age, gender, ethnicity, marital status, education, and income, background variables found important in previous research (see, e.g., Hilton, 1987; Grant & Harford, 1989; Caetano, 1990; Dawson, Grant & Harford, 1992).

This paper also has a methodological purpose. The external validity of alcohol abuse and dependence diagnoses in the general population is usually studied in a risk factor framework using logistic regression. In conventional risk factor analysis, a dichotomous diagnostic variable is regressed on a set of covariates or hypothesized antecedents. This paper suggests an alternative to logistic regression using a structural equation modeling framework. In structural equation modeling, one or more continuous latent variables that explain the covariation among a set of criteria are regressed on the covariates. In the case of multiple latent variables, different relationships with the covariates are possible. This alternative is suggested because it may offer more generality than logistic regression in understanding the external validity of diagnoses. Both logistic regression and structural equation modeling results will be presented in this paper.

Methods

Sample

The NHIS88 is a multi-stage complex sample household study of U.S. residents 18 years of age or older (see Massey, Moore, Parsons, Tadros, 1989). The alcohol-related questions were asked of one randomly selected individual in each household, resulting in a sample of 43,809. Of these individuals, 22,102 were classified as "current drinkers" by an affirmative answer to the question "In the past 12 months did you have at least 12 drinks of any kind of alcoholic beverage?" This paper will analyze data on the current drinkers. Given the large sample size, it is possible to split the sample into two random halves, using one half as a calibration sample to explore models and the other half for cross-validation of models suggested in the calibration stage.

The current drinkers were asked a set of symptom item questions with the introduction "In the past 12 months how many times have you..." To operationalize the 11 criteria of DSM-III-R and the proposed DSM-IV, 27 symptom items were used (see Grant, Chou, Pickering, Hasin, 1992; Grant, Harford, Hasin, Chou, Pickering, 1992). These symptom items are presented in Table 1 within their corresponding criteria. A diagnostic criterion was considered fulfilled if at least one of its symptoms was experienced at least twice in the last year. There were two exceptions. For the withdrawal criterion at least three positive symptoms were required, two of which additionally met this duration criterion. For the tolerance criterion, two positive symptoms were required, neither of which needed to meet the duration criterion. Each of the 11 criteria is a binary variable scored 0 or 1, with 1 denoting that the criterion is fulfilled.

Insert Table 1

The criteria will be related to a set of covariates consisting of alcohol consumption variables, family history of alcoholism, and demographic characteristics of the respondent. The variables chosen have been suggested as important predictors of alcohol-related problems in previous work (see, e.g., Hilton, 1987; Grant & Harford, 1989; Caetano, 1990; Dawson, Grant & Harford, 1992). Table 2 gives the definition of the variables and their means and standard deviations in the calibration sample of current drinkers. Of the 22,102 current drinkers, 18,650 individuals had complete data on all these covariates and were used in the analyses. The individuals with incomplete data did not differ in any important ways from those with complete data.

Insert Table 2

As shown in Table 2, three consumption variables are included: amount, frequency, and frequency of at least five drinks. These three variables appear to capture different aspects of drinking behavior. The correlations among the three variables are not as high as one might expect (the correlations are: .00 for amount-frequency, .43 for amount-frequency of at least five drinks, and .30 for frequency- frequency of at least five drinks). The family history of alcoholism variables are defined in line with Dawson, Grant & Harford (1992), using the responses to the two questions: "When you were growing up, that is, during your first 18 years, did you live with anyone who was a problem drinker or alcoholic?" and "Have any of your (other) blood relatives EVER been problem drinkers or alcoholics?"[1]. Familial transmissibility of alcoholism is evidenced by increasing effects when considering these dummy variables in the order Fh23 (family history positive for second or third degree relatives *only*), Fh1 (family history positive for first degree relatives *only*), Fh123 (family history positive for first *and* second or third degree relatives).

Statistical Analysis

The criteria will be related to the covariates using the statistical technique of structural equation modeling (see, e.g., Joreskog & Sorbom, 1979) generalized to dichotomous outcomes (see, e.g. Muthen, 1979, 1989) and implemented in the LISCOMP computer program (Muthen, 1987). The technical details of this approach will not be given here but the ideas of the modeling can be described as follows. The associations among the criteria are explained by a small number of underlying continuous factors sometimes referred to as latent variable constructs. In line with the analyses of Muthen et al. (1993), there are two such factors in the present case, one corresponding to less severe problems and the other corresponding to more severe problems. The notion is that these factors represent the fundamental dimensions being measured by the criteria and that the analysis should focus on these dimensions. In this case, a dichotomous diagnosis is not used. Instead, a "dimensional" representation is used where increasing factor values correspond to increased risk. This avoids the problem of choosing cutpoints and misclassification errors. A psychometric investigation in Muthén (1996) suggests that alcohol abuse and dependence diagnoses in general population surveys are likely to have low sensitivity, i.e. many of the cases are diagnosed as noncases. Cutpoints could, however, be chosen for the factors if classification of individuals is required.

The structural equation model used in this paper is shown in diagram form in Figure 1. It has three important sets of relationships: the relationships between the criteria and the factors (the measurement model), the relationships between the factors and the covariates (the structural regression equations), and the relationships between the criteria and the covariates (the direct effects).

Insert Figure 1

The relationships between the dichotomous criteria and the factors are described statistically by means of probit regressions. In the probit regression, the probability of a certain criterion being fulfilled is a function of the values of the factors as in latent trait theory and item response theory (for applications of this modeling idea to psychopathology, see, e.g. Duncan-Jones, Grayson, Moran, 1986; Muthen, 1989). The strength of these relationships gives information on how well the criteria serve as measurements of the factors.

The relationships between the factors and the covariates are described statistically by means of linear regressions. These regressions are termed structural equations in that they concern the more fundamental dependent variables of the factors, not their measurements. The main interest in the analysis is to estimate these structural equations.

The structural model states that the influence of covariates on criteria is mediated by the factors. In its most parsimonious form, there is no direct influence from any covariate to any criteria. This corresponds to the notion of a factor model for the criteria that is the same for all individuals in the sample. It is possible, however, for covariates to have direct effects on criteria. For example, if a dummy variable covariate, such as gender, has a direct influence on any criterion, the measurement parameters of the factor model are different for the two genders. The direct relationships between the criteria and the covariates are described by probit regressions.

The modeling strategy is to first apply the parsimonious structural model with no direct effects and then explore the need for including direct effects. In practice, the fit of the model to the data is usually improved significantly by the inclusion of a small number of direct effects. Often, these direct effects do not pose a serious threat to the validity of the factor model in that they are of relatively minor size. Direct effects can, however, reveal important subgroup differences. In this paper, the exploration of direct effects will be carried out in the calibration sample consisting of half of the 18,650 current drinkers with complete data on the variables used in the analysis. The remaining half, the validation sample, will be used to study how well the model produced in the calibration stage is reproduced in terms of fit and stability of estimated direct effects.

Conventional logistic regression results will also be presented for purposes of comparison. For the logistic regression analysis, the DSM-III-R definition of dependence will be used. Using this definition, individuals are classified as dependent if at least three criteria are fulfilled.

Interactions among covariates are included in both the structural equation and logistic regression models. Because of the large number of potential interactions among the set of covariates, the following selection process was used for determining which interactions to include in the structural equation model. First of all, the decision was made to include only two-way interactions. Then, a logistic regression was carried out for each of the 11 criteria and for DSM-III-R dependence. Interactions found to be significant for at least one of the 11 criteria or for the DSM-III-R analysis were included in the structural equation model. For the logistic regressions, a stepwise procedure was used to find significant two-way interactions.

There are two major reasons for including interaction effects in the structural equation and the logistic regression models. The first is related to the complex sample design used in NHIS88 which included an oversampling of blacks. Because both analyses are carried out without using sampling weights, there is a risk of distortions in estimates. This risk is minimized when oversampling factors such as ethnicity, and interactions involving them, are included among the background variables (see, e.g. Pfeffermann & LaVange, 1989). The second reason is that different subgroups of the population may be differentially at risk for alcohol problems given the same consumption behavior. For example, given the same alcohol intake, females may have stronger adverse effects. Again, the validation sample will be used to study how well the interaction effects produced in the calibration stage are replicated.

Results

As shown in the model fit summary of Table 3, the structural equation model with two factors makes a large improvement in fit over a unidimensional model in the calibration sample (the chi-square measure of model misfit drops from 2,049 with 424 degrees of freedom for the single-factor model to 1,138 with 376 degrees of freedom for the two-dimensional model and the drop is significant, $p < .001$; n=9,325). Including a small number of direct effects further improves the model. These direct effects are discussed after the main results. The measurement model estimates will be presented first, followed by the logistic regression estimates, the structural model estimates, and the direct effects estimates.

Insert Table 3

Measurement Model Estimates

The loadings of the measurement part of the model are shown in Table 4. The estimates show the same two-factor structure as found in Muthén et al. (1993). The first factor corresponds to less severe problems in that it is measured by criteria that are more prevalent, while the second factor corresponds to more severe problems. The less severe factor, interpreted as abuse, is particularly well measured by the criteria Larger and Hazard (see Table 1 for criteria definitions). The severe factor, interpreted as dependence, is particularly well measured by the criteria Cutdown, Giveup, Continue, and Relief. Table 4 shows that the proportions are considerably larger for the Larger/Hazard criteria as compared to the Cutdown/Giveup/Continue/Relief criteria. The correlation between the two factors is .68. The validation sample analysis replicates the measurement model estimates very closely.

Insert Table 4

Logistic Regression Estimates

The first row of Table 5 refers to the logistic regression analysis using the DSM-III-R dependence definition. The positive coefficients for the covariates of alcohol consumption, family history of alcoholism, being widowed/divorced/single or never married indicate that these variables are associated with increased probability of dependence, other factors being equal. Negative coefficients indicate protective factors so that the probability of dependence is decreased. This is the case for increasing age, being female, or being black, although the latter coefficient is not significant at the 1% level. These results are in line with previous research on alcoholism risk factors.

Structural Model Estimates

Table 5 shows the standardized parameter estimates[2] for the calibration sample structural model regressions for each of the two factors as well as the direct effects from the covariates to the criteria. Parameter estimates are not included for the validation sample although significance of these estimates is shown. Only estimates significant in both the calibration and validation samples will be discussed below.

The structural equation results are discussed first for the interaction effects among the covariates. There are eight interaction effects that are validated, but all effects are relatively small. Although the model including interaction effects fits the data better than the main effects only model, the significance and magnitude of main effects remains the same for the two models. To simplify the discussion, all main effects are discussed as though no interaction effects are significant.

Turning to the structural equation coefficients for the main effects, it should be noted that these are interpreted as partial regression coefficients just as in ordinary, linear multiple regression. Positive coefficients are seen in the rows for both the abuse factor and the dependence factor for the covariates of alcohol consumption, family history of alcoholism, being widowed/divorced/single or never married. This indicates that other things being equal, these variables are associated with elevated values of both factors and therefore increased probability of fulfilling each criterion. Increased age, being female and being black are associated with lower values of the factors, other things being equal. Given the variable standardization, the coefficient .17 for the dependence factor on Amount, for example, refers to a 17% standard deviation increase in the dependence factor when Amount increases by one standard deviation, other things being equal, while the coefficient .46 for Fh123 refers to a 46% standard deviation increase in the dependence factor when positive family history is present among first and second/third degree relatives as compared to when there is no type of positive family history, other things being equal.

Insert Table 5

Table 5 also shows important differences in how the two factors relate to the covariates as well as differences between the logistic regression and the structural modeling results. The coefficients of the structural regression for the dependence factor are significantly different from those of the abuse factor (a test of all coefficients being equal across the two factors obtained a chi-square value of 534 with 38 degrees of freedom, $p < .001$). It is seen that the variable Fh123, positive family history of alcoholism among first and second/third degree relatives, has a much stronger effect on the dependence factor than on the abuse factor. This fact cannot be revealed by the logistic regression since it considers a single dependent variable reflecting both the abuse and dependence factors. The family history coefficients also increase with increased genetic loading, going from Fh23 to Fh1 to Fh123. For the abuse factor the difference in these coefficients is, however, not significant at the 1% level (the test of equality of coefficients gives a chi square of 1.8 with two degrees of freedom, $p > .30$), but for the dependence factor it is strongly significant (the test of equality of coefficients gives a chi square of 12.0 with two degrees of freedom, $p < .01$), largely due to the big Fh123 effect.

Table 5 also shows a large differential effect on the two factors with respect to Age. While the differential family history effect is well replicated in the validation analysis, the differential age effect is, however, less pronounced in the validation (in the validation sample the abuse factor coefficient is -.34 and the dependence factor coefficient is -.20; both are significant at the 1% level).

Other differences in coefficients for the two factors occur for Female, Black, Hispanic, and Education, but it is only for Hispanic and Education that the difference is clearly replicated in the validation analysis. For Hispanic, Table 5 shows that the coefficients obtain opposite signs in the abuse factor and dependence factor regressions. It is interesting to note that the logistic regression for dependence does not show a significant effect for Hispanic in the calibration sample or in the validation sample. The structural model estimates of Table 5 suggest that this is because the effects on the two factors cancel each other out. Furthermore, Education has a negative, protective, effect only on the dependence factor, not on the abuse factor.

An R^2 quantity, describing amount of variation explained by the regression, can be computed for the structural equations just as in ordinary regression. The abuse factor has a rather large amount of its variation explained by this set of covariates, 52%, while only 21% of the dependence factor variation is explained.

Direct Effect Estimates

The exploration of direct effects in the structural model led to the inclusion of effects that provided important model improvement (the chi square value of model misfit dropped significantly from 1,138 to 774 with a drop of only 21 degrees of freedom, $p < .001$). As indicated in Table 5, a direct effect was found for the variable Female showing a negative direct effect on the criterion Hazard, recurrent drinking in situations in which it is physically hazardous. This effect should be viewed in conjunction with the structural estimates just discussed. The variable Female has a negative coefficient for both the abuse factor and the dependence factor so that females have decreased levels of the factor scores relative to males and therefore have lower probability of fulfilling all the criteria, other things being equal. The negative direct effect states that relative to males, females have an additional decrease in the probability of fulfilling the Hazard criterion. Table 4 shows that the Hazard criterion measures the abuse factor. Since this additional decrease is not a function of the factors, it implies that the Hazard criterion does not measure abuse in the same way for males and females; the criterion is relatively more benign for males. This direct effect was well-replicated in the validation sample analysis. Other direct effects, such as the negative effects of Age on Majorole and Wds on Withdrawal and the positive effect of Poverty on Tolerance, were also well-replicated.

Discussion

One finding of this study is the validation of the two-dimensional model for alcohol abuse and dependence found by Muthén et al. (1993) using a general population sample. Relating the DSM-III-R and DSM-IV criteria to background variables, the same model was recovered. The first factor corresponds to less severe problems than the second factor as indicated by the level of prevalence for the criteria that measure each factor well. The factors may be interpreted as representing an abuse dimension and a separate, although highly correlated, dependence dimension. By relating the criteria to a set of covariates, the model fit statistics for the structural equation analysis give additional support for the need for two dimensions instead of one to represent covariation among the criteria. The need for two dimensions is further evidenced by the two factors having different relationships to the background variables.

It is interesting to note that the model departs from diagnostic schemes related to the DSM-III-R and DSM-IV. In DSM-III-R, an individual is diagnosed as alcohol dependent if at least three out of nine criteria are fulfilled and some symptoms fulfill the duration criterion of having persisted for at least one month or occurred repeatedly over a longer period of time. Such a definition assumes that all criteria reflect a single dimension along which dependence becomes more severe. The notion of unidimensionality is also clear from the DSM-III-R and DSM-IV severity modifiers of mild, moderate, severe that are defined by the number of dependence criteria fulfilled. Also, any three of the criteria can be fulfilled for dependence diagnosis, suggesting that all criteria are equally good measurements of the single dimension. The two-dimensional model of Muthén et al. (1993) and replicated here presents a different view of how the criteria measure abuse and dependence. The more severe alcohol-related problems of the dependence factor and the less severe problems of the abuse factor do not define opposite ends of the same dimension, but define phenomena of a distinct kind. The DSM-III-R dependence definition that includes the abuse criteria blurs the distinction between the two disorders. The dependence definition is too broad and inclusive, containing criteria that are quite common in a nationally representative sample, and therefore leads to an overestimation of the prevalence of dependence.

A second important set of findings concerns assessing risk factors of alcohol abuse and dependence. In this study, the importance of risk factors is determined by the relationships between the covariates and the factors. The differences found in these relationships indicate that the set of important covariates, the risk factors, for the two factors differ.

In terms of risk factors, one of the most striking findings relates to the increased risk of positive family history of alcoholism among both first and second or third degree relatives. Positive family history of this kind has a much stronger effect on the dependence factor than on the abuse factor. The effect on the dependence factor corresponds to about 46% of its standard deviation, while the abuse factor effect is less than half of that. The validation sample analysis replicates this finding. The increased risk is observed despite the fact that other risk factors are controlled for in the analysis, including measures of alcohol consumption and sociodemographic characteristics of the individual. Family history of alcoholism may also exert an influence on alcohol problems indirectly via increased consumption, but the effects observed in these analyses refer to direct effects, over and above alcohol consumption. The findings may be taken as an indication that the dependence factor is the more important factor to consider in terms of alcohol problems and that the criteria that measure the dependence factor well should be emphasized in diagnoses of alcohol dependence. The degree of familial alcoholism also appears to be of importance. The results indicate that having positive family history among first degree relatives only and in addition to second or third degree relatives, respectively, gives increasingly large positive effects on the factors, in particular the dependence factor.

Another interesting finding in relation to risk factors is that the two factors have different relationships with the covariate Hispanic. These effects are given relative to the white, non-Hispanic subgroup. For the abuse factor the relationship is negative. For the dependence factor, however, it is positive suggesting that Hispanics are more at risk for alcohol dependence than white, non-Hispanic individuals, other things being equal. The logistic regression using the DSM-III-R dependence definition overlooks this risk factor. The structural model effects for the two factors are about 20% of a standard deviation of the two factors and this magnitude is replicated in the validation sample. This is an important new finding made possible by the large sample size of NHIS88 and the separation of the set of criteria into two factors. The difference in risk for abuse versus dependence in Hispanics warrants further research.

Education also shows differential effect on the two factors. While the effect of on the abuse factor is negligible, the protective effect for the dependence factor of a standard deviation increase in education corresponds to 10-12% of the dependence factor standard deviation in the calibration and validation samples.

Interactions among covariates proved to be of little importance. Few effects were of practical importance and none of practical importance was replicated in the validation analysis.

The set of covariates used in this study explain different amounts of the variability for the two factors. The abuse factor has a rather large amount of its variation, 52%, explained by the set of covariates. In contrast, the 21% variation explained for the dependence factor shows that more specific risk factors need to be sought for these more severe problems.

The importance of measurement invariance of criteria across subgroups of the population is investigated in this study through the direct effects of the covariates on the criteria. Several criteria showed subgroup differences. The important ones are Hazard showing a gender difference, Majorole showing an age difference, Withdrawal showing a difference for Widowed/divorced/separated, and Tolerance showing a difference for Never married and Poverty. This indicates that some criteria are not invariant across certain groups. Given how the Hazard criterion is operationalized in the NHIS88 (see Table 1), the lower female probability to fulfill the Hazard criterion may be due to females being more likely than males to be a passenger rather than a driver of a car after having had too much drink, reducing this risk of hurting themselves or others. The lower probability at higher age to fulfill the Majorole criterion may be due to major role obligations being more critical at a higher age. The reasons for the decrease in the probability of fulfilling the Withdrawal criterion for those widowed, divorced, or separated and for fulfilling the Tolerance criterion for those never married, as well as the increase in the probability to fulfill the Tolerance criterion for those living in poverty, need to be further researched. These effects represent non-invariance of the measurements which can distort risk factor assessment and group comparisons in prevalence estimates.

In conclusion, the analyses show that the use of the DSM-III-R dependence definition does not provide as clear a picture of risk factors as obtained via the structural model. Relative to the structural model, important relationships are confounded in the logistic regression for the covariates of family history of alcoholism and Hispanic. The family history finding supports the suggestion of Muthén et al. (1993) that the criteria of the dependence factor are those most important for dependence diagnosis. The measurement non-invariance of several of the criteria across subgroups of individuals implies that the DSM-III-R definition of dependence may be distorted. This gives rise to bias in risk factor assessment and in comparisons of prevalence estimates for these subgroups.

The structural equation modeling approach used in the study has several methodological strengths in terms of understanding the external validity of diagnoses and assessing risk factors. One important strength is that structural equation modeling allows the criteria to be represented by more than one dimension. In this way, the different dimensions are allowed to have different risk factors. A second strength is that measurement non-invariance of criteria across subgroups of individuals in the population can be detected and allowed for, so that the risk factor assessment is not distorted by such non-invariance. A third strength is that the dependent variables are the continuous factors and not a dichotomous diagnosis variable. This avoids the problem of choosing cutpoints and misclassification errors. A psychometric investigation in Muthén (1996) suggests that alcohol abuse and dependence diagnoses in general population surveys are likely to have low sensitivity, i.e. many of the cases are diagnosed as noncases. A fourth strength is that maximum information in the criteria is retained because responses to the set of criteria are not collapsed into an aggregate sum as in the DSM-III-R definition of dependence, but the multivariate responses are maintained. In sum, the structural equation modeling approach provides a general and powerful technique for validity and risk factor assessment in psychiatric epidemiology.

Footnotes

1 I thank Dr. Dawson and Dr. Grant at NIAAA for providing the data on these variables.

2 The structural estimates are given in standardized form, corresponding to dependent variable factor variances of one and standardized continuous background variables. Standardization has not been carried out for background variables which are 0/1-coded dummy variables. The logistic regression estimates have been standardized with respect to background variables in the same way. Since the logistic regression coefficients are given in the customary logit scale they are not comparable to the structural regression coefficients, but only sign and significance of estimates can be compared.

References

American Psychiatric Association (1987). Diagnostic and Statistical Manual of Mental Disorders, 3rd edition, revised (DSM-III-R) Washington DC, American Psychiatric Association.

American Psychiatric Association (1992). DSM-IV Options Book. Washington, DC, American Psychiatric Association.

Caetano, R. (1990). Correlates of DSM-III-R alcohol dependence in treatment and general populations. Drug and Alcohol Dependence.

Dawson, .D. A., Harford, T. C., Grant, B. (1992). Family history as a predictor of alcohol dependence. Submitted for publication.

Duncan-Jones, P., Grayson, D. A., & Moran, P. A. P. (1986). The utility of latent trait models in psychiatric epidemiology. Psychological Medicine, 16, 391-405.

Grant, B. F., Chou, P. , Pickering R. P. , & Hasin, D. S. (1992). Empirical subtypes of DSM-III-R alcohol dependence: United States, 1988. Drug and Alcohol Dependence, 30, 75-84.

Grant, B. F., & Harford, T. C. (1989). The relationship between ethanol intake and DSM-III alcohol use disorders: A cross-perspective analysis. Journal of Substance Abuse, 1, 231-252.

Grant, B. F., Harford, T. C., Hasin, D. S., Chou, P., & Pickering, R. (1992). DSM-III-R and the proposed DSM-IV alcohol use disorders, United States 1988: A nosological comparison. Alcoholism: Clinical and Experimental Research, 16, 215-221.

Hilton, M. E. (1987). Demographic characteristics and the frequency of heavy drinking as predictors of self-reported drinking problems. British Journal of Addiction, 82, 913-925.

Joreskog, K. G. & Sorbom, D. (1979). Advances in Factor Analysis and Structural Equation Models. Cambridge Mass.: Abt Books.

Massey, J. T., Moore, T. F., Parsons, V. L. & Tadors, W. (1989). Design and estimation for the National Health Interview Survey, 1985-1994. Vital and Health Statistics, 2, 1-33.

Muthen, B. (1979). A structural probit model with latent variables. Journal of American Statistical Association, 74, 807-811.

Muthen, B. (1987). LISCOMP. Analysis of Linear Sructural Equations with a Comprehensive Measurement Model. Theoretical integration and user's guide. Mooresville, IN: Scientific Software.

Muthen, B. (1989). Dichotomous factor analysis of symptom data. In Eaton and Bohrnstedt (eds.), Latent Variable Models for Dichotomous Outcomes: Analysis of Data from the Epidemiological Catchment Area Program. Sociological Methods & Research, 18, 19-65.

Muthen, B. (1996). Psychometric evaluation of measurement properties: Application to a two-dimensional model of alcohol abuse and dependence. Drug and Alcohol Dependence, 41, 101-112.

Muthen, B., Grant, B., & Hasin, D. (1993). The dimensionality of alcohol abuse and dependence: Factor analysis of DSM-III-R and proposed DSM-IV criteria in the 1988 National Health Interview Survey. Addiction, 88, 1079-1090.

Pfeffermann, D. & LaVange, L. (1989). Regression models for stratified multi-stage cluster samples. In Skinner, Holt, & Smith (eds.), Analysis of Complex Samples (pp. 237-260). New York: Wiley & Sons.

Table 1
Diagnostic Criteria for DSM-III-R and DSM-IV Alcohol Abuse and Dependence and Associated Questionnaire Items

Analysis Acronym	Diagnostic Criterion	Questionnaire Items	Proportion*
LARGER	Drinking in larger amounts or over a longer period than the person intended.	• Ended up drinking much more than you intended to.	.25
		• Found it difficult to stop once you started.	.08
		• Kept on drinking for a longer period of time than you intended to.	.14
CUTDOWN	Persistent desire or one or more unsuccessful efforts to cut down or control drinking.	• Tried to cut down or stop drinking and found you couldn't do it.	.01
		• Wanted to cut down or stop drinking and found you couldn't do it.	.01
TIMESPENT	Spent a great deal of time obtaining alcohol, drinking, or recovering from drinking.	• Spent a lot of time drinking or getting over the effects of drinking.	.03
MAJOROLE	Frequent intoxication or withdrawal symptoms when expected to fulfill major role obligations at work, school, or home.	• Stayed away from work or gone to work late because of drinking or from a hangover.	.03
		• Gotten drunk instead of doing the things you were supposed to do.	.04
		• Been so hung over that it interfered with doing things you were supposed to do.	.04
HAZARD	Recurrent drinking in situations in which it is physically hazardous.	• Driven a car after having too much to drink.	.10
		• Done things when drinking that could have caused you to be hurt.	.07
		• Done things when drinking that could have caused someone else to be hurt.	.04

(table continues)

* Proportion in the calibration sample of current drinkers (n=9,325) admitting to having had this happen at least twice in the last 12 months.

Table 1 (continued)

Analysis Acronym	Diagnostic Criterion	Questionnaire Items	Proportion*
GIVEUP	Important social, occupational, or recreational activities given up or reduced because of drinking.	• Given up or cut down on activities or interests like sports or associations with friends, in order to drink.	.01
		• Lost ties with or drifted apart from a family member or friend because of your drinking.	.01
CONTINUE	Continued to drink despite knowledge of a persistent or recurrent social, psychological, or physical problem that is caused or exacerbated by drinking.	• Continued to drink alcohol even though it was a threat to your health.	.02
		• Kept drinking even though it caused you emotional problems.	.02
		• Kept drinking even though it caused you problems at home, work, or school.	.01
		• Had a spouse or someone you lived with threaten to leave you because of your drinking.	.01
TOLERANCE	Tolerance.	• Found that the same amount of alcohol had less effect than before.	.09
		• Found that you had to drink more than you once did to get the same effect.	.04

(table continues)

* Proportion in the calibration sample of current drinkers (n=9,325) admitting to having had this happen at least twice in the last 12 months.

Table 1 (continued)

Analysis Acronym	Diagnostic Criterion	Questionnaire Items	Proportion*
WITHDRAWAL	Characteristic withdrawal symptoms.	• Been sick or vomited after drinking, or the morning after.	.09
		• Felt depressed, irritable, or nervous after drinking or the morning after.	.10
		• Heard or seen things that weren't really there after drinking, or the morning after.	.01
		• Found yourself sweating heavily or shaking after drinking, or the morning after.	.03
RELIEF	Drinking to relieve or avoid withdrawal symptoms.	• Taken a drink to keep yourself from shaking or feeling sick either after drinking or the morning after.	.01
LEGAL	Recurrent alcohol-related legal or interpersonal problems.	• Been arrested or had trouble with the police because of your drinking.	.01

* Proportion in the calibration sample of current drinkers (n=9,325) admitting to having had this happen at least twice in the last 12 months.

Table 2

Descriptive Statistics for Current Drinkers in the Calibration Sample (n=9,325)

Variables	Description	Mean	S.D.
Amount	Number of drinks per day on average	2.73	2.22
Freq	Frequency of drinking (days per week)	2.16	2.31
Freq5+	Number of days per week had 5 or more drinks	0.27	0.89
Fam0	No family history of alcoholism	0.59	0.49
Fh23	Positive family history for second/third degree relatives only	0.17	0.37
Fh1	Positive family history for first degree relatives only	0.15	0.36
Fh123	Positive family history for first and second/third degree relatives	0.09	0.28
Age		40.58	16.16
Female		0.48	0.50
Black		0.10	0.30
White		0.85	0.35
Hispanic		0.05	0.21
Wds	Widowed, divorced, or separated	0.21	0.41
Nmar	Never married	0.24	0.43
Education	Education is scored as 1-4 for the categories 0-11 years. 12 years (high school graduate), some college, college graduate	1.61	1.01
Poverty	The NHIS88 Poverty Index	0.09	0.29
Misspov	Dummy variable for missing data on poverty	0.04	0.20

Table 3

Structural Equation Model Fit Summary
(n=9325)

	Chi-Square	Degrees of Freedom
One-Factor Model	2049	424
Two-Factor Model	1138	376
Two-Factor Model with 21 direct effects	774	355

Table 4

Standardized Parameter Estimates for the Structural Equation and Logistic Regression Models for the Calibration Sample (n=9,325)

	Factor Loadings		
	Abuse	Dependence	Proportions
Larger	0.77--	0.00--	0.28
Cutdown	0.00--	0.73--	0.02
Time	0.36*†	0.45*†	0.03
Majorole	0.44*†	0.35*†	0.07
Hazard	0.62*†	0.17*†	0.12
Giveup	0.04	0.72*†	0.01
Continue	0.10 †	0.63*†	0.04
Toler	0.20*†	0.47*†	0.04
Withdraw	0.21*†	0.56*†	0.02
Relief	-0.04	0.75*†	0.01
Legal	0.05	0.65*†	0.01

Table 5
Structural Regression Coefficients

	Amount	Freq	Freq5+	Fh23	Fh1	Fh123	Age	Female
Logistic Regression of DSM-III-R Dependence								
	0.56*†	0.43*†	0.21*†	0.29*†	0.46*†	0.65*†	-0.62*†	-0.66*†
Abuse	0.28*†	0.25*†	0.10*†	0.12*†	0.15*†	0.19*†	-0.46*†	-0.10*
Dependence	0.17*†	0.14*†	0.12*†	0.16*†	0.28*†	0.46*†	-0.06 †	-0.22*
Larger							0.12*	
Cutdown	-0.01							
Time								
Majorole	-0.02*					0.13*	-0.12*†	
Hazard								-0.15*†
Giveup			-0.04*					
Continue								
Toler							-0.03 †	-0.08
Withdraw		-0.06*†						0.07
Relief								
Legal								

Table 5 continued

	Black	Hisp	Wds	Nmar	Educ	Poverty	Misspov	FemXBla [1]
Logistic Regression of DSM-III-R Dependence								
	-0.32	0.06	0.63*†	0.57*†	-0.07 †	0.14	-0.19	
Abuse	-0.17*	-0.20*†	0.22*†	0.22*†	0.01	0.00	-0.21*	0.08
Dependence	-0.15	0.19 †	0.17*†	0.21*	-0.10*†	0.02	-0.24	0.02
Larger								
Cutdown		0.14						
Time								
Majorole								
Hazard								
Giveup								
Continue								
Toler				-0.13 †		0.23*†		
Withdraw			-0.11*†					
Relief								
Legal								

	FemXAge	FemXAmt	FemXFq5	BlaXAge	BlaXWds	BlaXAmt	HisXPov	HisXAmt
Logistic Regression of DSM-III-R Dependence								
	-0.43*		0.27*†			-0.33*		
Abuse	-0.04	0.06*	0.04*†	0.05*†	-0.10	-0.03*	0.37	-0.03*
Dependence	-0.00	-0.02	0.03	-0.02	0.41*	-0.02	0.06	0.02
Larger								
Cutdown								-0.02*
Time								
Majorole								
Hazard								
Giveup								
Continue								
Toler	-0.09							
Withdraw								
Relief								
Legal								

Table 5 continued

	F23XAmt	F1XPov	F123XAg	F123XFq	F123XEd	WdsXFq	NmaXAge	EduXAmt
Logistic Regression of DSM-III-R Dependence								
		0.72*		0.32*				0.13*†
Abuse	-0.01	-0.09	0.01	0.02	-0.04*†	-0.01	-0.03 †	0.07*†
Dependence	0.01	0.14	-0.01	0.05*	-0.02	0.03 †	-0.03	0.01 †
Larger								
Cutdown								
Time								
Majorole	0.03*		0.04*					
Hazard								
Giveup								
Continue								
Toler							-0.13*†	
Withdraw								
Relief								
Legal								

	EduXFq5	PovXAge	PovXAmt	AgeXAmt	AgeXFq	AgeXFq5
Logistic Regression of DSM-III-R Dependence						
					-0.18*	
Abuse	0.03*†	-0.01	-0.03*	0.05*†	-0.05*†	-0.05*†
Dependence	0.07*†	0.03	0.05*†	0.07*†	-0.03	-0.02
Larger						
Cutdown						
Time						
Majorole						
Hazard						
Giveup						
Continue						
Toler						
Withdraw						
Relief						
Legal						

-- Parameter is fixed to set the metric of the factors.
* Significant at the .01 level in the calibration sample.
† Significant at the .01 level in the validation sample and of the same sign as in the calibration sample.
[1] Interaction variables are denoted by "X." For example, FemXBla stands for the Female and Black interaction. In the interactions, the Fam variables have been abbreviated as F and the frequency variables as Fq.

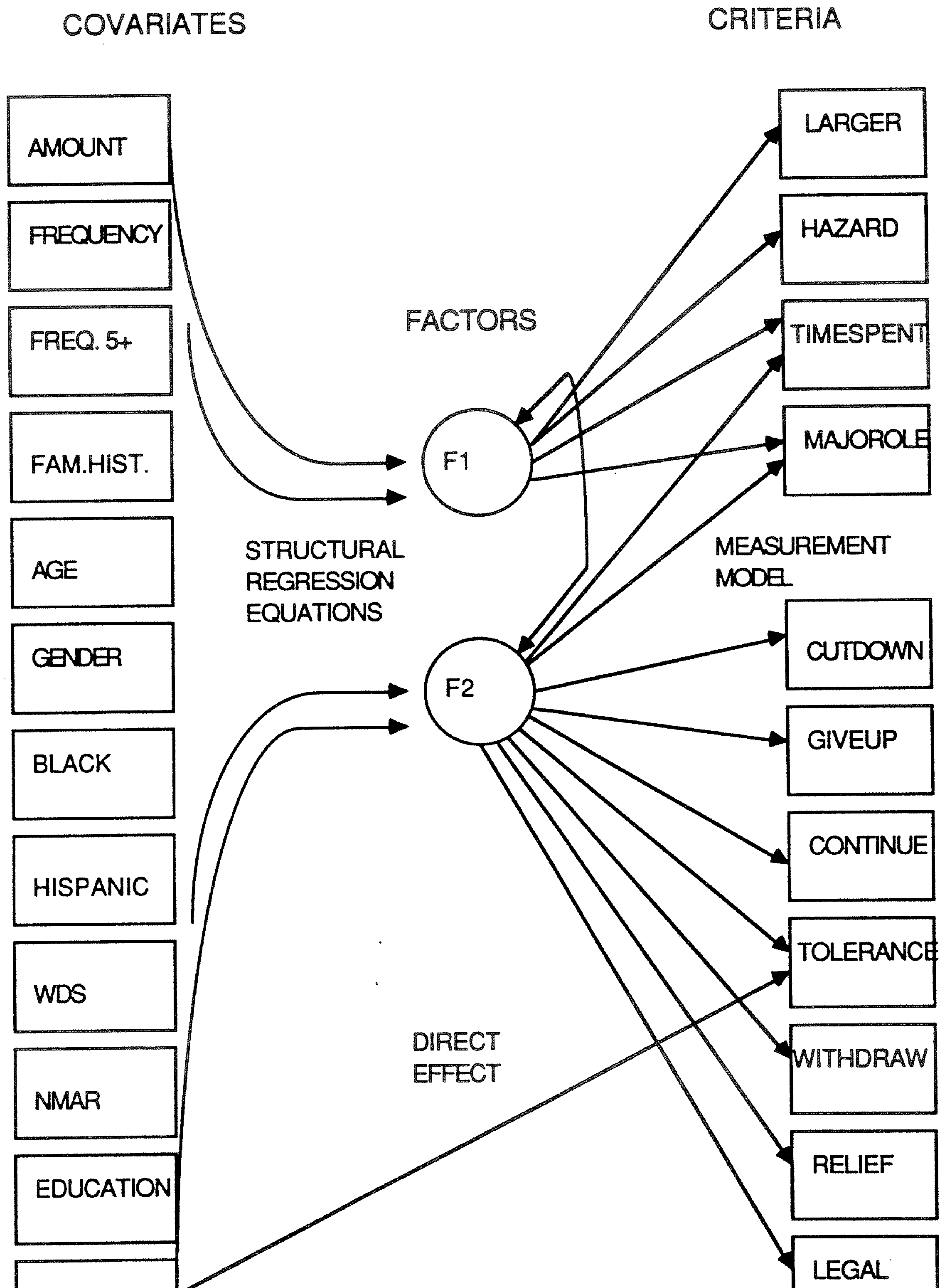

FIGURE 1: STRUCTURAL EQUATION MODEL